

Towards a European university system

Soon, the European Commission's writ should run to higher education. Present arguments in France about the constitution of universities point to the need for a coherent European policy in the field.

ONE of the paradoxes of contemporary Europe is that, while "ever closer union" is happening, if fitfully, nobody quite knows what will become of Europe's universities. That is a strange oversight, given the value that Europeans, like others, attach to higher education not merely as a means of educating their young but also as a source of culture, scholarship and innovation. But the oversight must be deliberate. The Treaty of Rome, 40 years ago, said nothing about universities. The European Single Act of 1986 mentioned research, but not higher education. The Maastricht Treaty, yet to be ratified, will enable the European Commission to take an interest in educational affairs, but only within terms of reference defined from time to time by the member states of the European Communities (EC). Meanwhile, member governments set about the reform of their own systems in isolation from each other and sometimes, to judge from the report from France on page 5 of this issue, by reaction against what they think has happened elsewhere.

The continent that invented the idea of the university can surely do better than that. The case of the French universities is a good illustration of the need. Republican tradition had established the right to equal access long before the traumatic upheavals of 1968 reminded the authorities that students are not mere ciphers. Since then, the French system has been securely launched towards self-betterment. The funding of the system has been generous; the appearance of Napoleonic centrism notwithstanding, universities have grown to differ from each other under the pressure of serving their students' needs; and scholarly ambitions have been greatly magnified. Left to itself, the French system would probably evolve into a pattern as diverse as any. The present argument about the formal devolution of authority from the centre to individual universities sounds specious.

That is not entirely the case. Countries with centralized university systems derive some benefits from the arrangement, France more than most. Uniform conditions of entry can be defined, teaching standards can be maintained by external invigilation and the merits of qualifications can be determined with a degree of objectivity not otherwise attainable. France has given the rest of the world a lesson by its arrangements for prequalifying entrants to research degree courses, which avoid much waste (of people as well as money). But there are drawbacks to centralism for all. One is bureaucracy. Another is the difficulty of legislating for outstanding universities within a supposedly uniform system; Japan, for example, will not have an outstanding

university until the twin constraints of its ministries of education and of finance are loosened.

The lesson to be learned from that may seem a weasel compromise, but is good sense; national university systems need to be a mixture of differently constituted institutions, some free to follow elitist paths, others more tightly constrained by what their paymasters believe are the benefits to be derived from their continued existence. For what it is worth, even France has a good approximation of such a system: what are the *grandes écoles* but frankly elitist institutions? Republicans delude themselves if they believe otherwise simply because Napoleon had a hand in founding them.

That points to a role for the European Commission in fostering higher education in Europe. Starting from the position that Europe's national governments will not readily or quickly give up their fondness for managing universities (but in Germany, responsibility rests with the *Länder* governments, as in Switzerland with the cantons), the Commission could usefully set out to foster (and even partially to finance) those institutions already so independent and international in outlook that they could function as if they were already European universities. It would not be unreasonable to require, in return, freedom of access by all European students and the recruitment of teachers from all over.

Those of despondent habit will say that it would be politically impracticable to set out on such a course, given the general wish of governments to get back more from the common enterprises to which they belong than they contribute to them. But that is not necessarily the case. Some would even find it helpful with the endless problem of marrying the pursuit of egalitarianism and elitism. Others, of generous disposition, would benefit from knowing that some at least of their students had found their way to more excellent places than there were at home. Still others would find it convenient to have models their own institutions and academics could emulate. If the Commission played its cards tactfully, it could successfully launch a scheme like this. The greatest difficulty would be to avoid replacing control by national government by control from Brussels.

The case for following this course, with all its risk, is broader. Everywhere in Europe, the cost of higher education is an increasing embarrassment, partly because of enlarged demand, and the quality of even outstanding institutions may be eroded by the pressures of chronic penny-pinching, a recipe for making sure that Europe will lack universities that match its own ambitions. That would be a great misfortune. □



European patent for PCR enzyme clouded by Russian claim

London. A Russian scientist whose research group was one of the first to isolate a thermostable enzyme from the bacterium *Thermus aquaticus* is claiming that his enzyme is identical to the widely used *Taq* DNA polymerase for which the US biotechnology company Cetus Corporation subsequently received a US patent.

Stanislav Gorodetsky, now head of the laboratory of molecular genetics at the Institute of Gene Biology in Moscow, says that the figure of 60,000 daltons identified as the enzyme's molecular weight in the original paper describing its purification was a "misprint". The printed figure has been used by Cetus to back its claim that the two enzymes are different; the corrected figure, says Gorodetsky, suggests they are identical.

Gorodetsky's claim adds a new twist to the bid by the Swiss-owned pharmaceutical company Hoffman-La Roche to secure European patent rights on *Taq* polymerase. The enzyme is used worldwide in the polymerase chain reaction (PCR) process, for which Roche bought the patent rights (including the rights to *Taq* polymerase) from Cetus in 1991 for \$300 million; it is currently marketed for PCR use exclusively through Perkin-Elmer Corporation, and has annual sales estimated last year at \$25 million in Europe alone.

The patent on *Taq* polymerase has been controversial, partly because the enzyme is still seen by many scientists as a "natural" product. Thomas Gold, for example, the microbiologist who originally discovered the bacterium during a visit to hot springs in Yellowstone Park in 1965, says he considers the organism to be in the public domain.

But arguments about the equitability of the *Taq* patent cut little ice with patent offices, since their concern has been primarily on questions of novelty and usefulness. The main focus of disagreement has been whether Cetus was justified in claiming that its scientists were first to purify the enzyme, given the acknowledged existence of previous reports — particularly that published by Gorodetsky and his colleagues, based on the work of postgraduate student A. Kaledin (*Biokhimiya* 45, 644–651; April 1980).

Kary Mullis, the inventor of PCR, has admitted that it was the Russian paper that first gave him the idea of using a polymerase obtained from the *T. aquaticus* bacterium as the thermostable enzyme needed for continuous re-use in the successive heating and cooling steps that amplify DNA fragments. (The previous enzyme, Klenow polymerase, had to be replenished after every step.)

Despite the Russian publication, Cetus was able to persuade the US Patent Office of the novelty of its own *Taq*, purified from the

bacterium by researchers David Gelfand and Susannah Stoffel; the US patent was awarded in December 1989. But last summer, an examiner at the European Patent Office in Munich, issuing a preliminary response to a patent application from Roche, gave the opinion that the Russian paper contradicted the novelty of the claims being made by Roche, since it appeared to refer to the same enzyme.

Roche has challenged this interpretation. In a revised patent submission, filed at the end of February, Roche defends its claim by saying that several technical features of the enzyme described in the Russian paper are different from those in its own submission. In particular, Roche points out that the molecular weight of the Russian enzyme, measured by two separate techniques, was described by Gorodetsky and his colleagues as being either 62,000 or 60,000 daltons. The enzyme purified by Cetus has a molecular weight of 86,000 to 90,000 daltons.

However Gorodetsky has explained that the figure of 60,000, based on a comparison of the effects of centrifugation of polymerase sediments with those of bovine serum albumin, is a misprint. The actual value calculated, he says, was about 80,000 — much closer to the Cetus figure.

Referring to the alternative estimate of 62,000 daltons, obtained using electrophoresis on a polyacrylamide gel, Gorodetsky says that its accuracy is questionable, as the polymerase and the markers used for comparison were put in different tubes.

The impact of Gorodetsky's claims will be limited by the fact that no correction was published (a follow-up paper by the same authors was rejected by another publication). Moreover, Roche still has a significant claim to the rights on its genetically engineered recombinant *Taq* polymerase.

As well as confusing the European patent application, the Russian claim, if con-

firmed, could strengthen the hand of the US biochemical supply company Promega Corporation, which is challenging the US patent in New Jersey courts (see *Nature* 362, 686; 1993) — a move widely seen as the company's reaction to a suit filed against it by Roche alleging infringement of a licensing agreement between the companies.

Randall Dimond, chief technology officer at Promega, says the company's own scientists have shown that the Gorodetsky and Cetus enzymes may be identical, and that the Russian paper contains "an error in the understanding of molecular weight".

Despite the continuing controversy, Roche says it is confident that its revised patent submission now meets all the novelty and inventiveness requirements of the European Patent Convention. It has pointed out, for example, that the difference in published molecular weights was not the only reason that Cetus believed it had purified a separate enzyme; other factors include a higher level of activity, and a greater ability to maintain this ability in moving from an alkali to an acidic environment.

Claude Montandon, Roche's director of licensing and new technology assessment, says the company is optimistic that it will be allowed the *Taq* patent "in the near future". He adds that, together with Perkin-Elmer, Roche is in the process of fine-tuning a licensing policy covering its DNA polymerase enzymes, although it is not at present prepared to provide any further details.

Meanwhile, Gorodetsky says that the Russian team's discovery was not patented when the enzyme was isolated in 1978 because the only way of doing so at that time would have been through an inventor's certificate — which would have practically ruled out the possibility of publishing the data. "We felt that it was more important to bring the results to public notice," he says.

David Dickson

Brain drain smaller than feared

Moscow. Only a few thousand researchers have left Russia permanently since the dissolution of the Soviet Union, according to a new report by the Russian Academy of Sciences. Although that represents less than 1 per cent of the estimated scientific work force, it does not measure the quality of those who have emigrated.

The academy, in an annual report by its analytic centre, paints a dismal picture of the future of Russian science and says that hope for an improved economic situation has all but vanished. But it also criticizes the government and the scientific community for failing to adopt reforms that could improve conditions in the short term, including setting priorities among various disciplines to make the best use of scarce resources.

The study says that not even the current request before the Russian Parliament for 673 billion rubles for science is likely to help without "the introduction of non-traditional methods and means of financing science". The report also emphasizes the role of international contributions.

Vladimir Pokrovsky

Can Europe put Columbus together again?

Brussels. A week is an eternity in space politics. Bloodied by US vacillation over an international space station, the European Space Agency (ESA) heaved a sigh of relief last week when President Bill Clinton affirmed his commitment to an international modular station based on the original Freedom design.

"We couldn't have hoped for a better outcome," says Lanfranco Emiliani, head of ESA's Columbus programme. ESA had panicked earlier this month when it became known that none of NASA's three options met the \$9 billion budgetary ceiling demanded by the US president (see *Nature* 363, 569; 1993). The agency feared that Clinton might kill the space station, wasting the millions of dollars it has already spent on Columbus, or even opt for the cheapest "big tin-can" redesign, thus making ESA's attached module redundant. Pulling out seemed on the cards.

Clinton's unexpected proposal for a space station combining elements of options A (a simpler version of Freedom) and B (a truncated version of Freedom, which can be expanded as budgets allow) has allayed ESA's fears. Although the fudge obscures what is being axed and what is being kept, it promises to cost less than the alternatives, and is expected to receive the support of Congress.

ESA is also pleased that Clinton post-

poned a decision to increase the inclination of the orbit of the space station from 28.8° to 51.6° to allow the station to be reached from Russian launch sites. ESA had protested that to launch Columbus into this higher orbit would require using the Russian Proton rocket, a competitor of Ariane. NASA is apparently considering a compromise orbit, which the Russians could still reach, although only by flying over China.

Clinton's decision has restored confidence in Europe, where support for the station has crumbled over the past few months. "There is now the political will to continue with the space station," says Klaus Berge, director of the German programme. Whether there is a financial way is another matter. Italy has said it cannot afford to meet the commitments it made in Granada last November (see *Nature* 360, 199; 1992), while ESA itself has reserved its position on continued participation in the light of "Europe's economic difficulties". ESA is relying on further reducing the costs of Columbus to avert internal strife, says Berge.

Continued support for the station in Europe is likely to be forthcoming only if the present *ad hoc* collaboration is transformed into a truly international project. The stumbling block to a *ménage à trois* between Europe, Russia and the United States is that all three are competitors in the commercial satellite launch market (see sidebar). And

Launch wars

As Europe, Russia and the United States discuss prospects for greater cooperation in an international space station, all three are engaged in a battle for a share of the commercial satellite launch market.

The United States and Europe are concerned that, unless Russia's entry is regulated, it could destabilize the market through dumping; Russia can undercut by half the \$62 million cost of an Ariane launch. But Europe and the United States have failed to agree on acceptable terms, and are already divided over the rules of fair play: the United States accuses Europe of subsidizing Arianespace, while ESA wants the United States to open government contracts to foreign competition.

Independently, the United States this month agreed tariffs and quotas with Russia for the launch of US satellites using its Proton rocket. But Europe immediately accused the United States of manipulating events to destabilize Ariane as the market leader while protecting its own launchers. ESA has since made a similar agreement with the Russians, but it remains to be seen how long Russia will be content to restrict its exports to the agreed quotas. **D.B.**

while Europe and the United States in public make much of their efforts to "save" Russian space competence, officials privately admit that both are jealously competing for access to Russian technology. **Declan Butler**

Monoglot filing urged for European patents

Munich. The Munich-based European Patent Office (EPO) last week announced a 5 per cent increase in the number of patent applications in 1992, after 1991 saw the first drop in its 15-year history. Half of the 58,000 applications were from member states, with Germany topping the list with 19.5 per cent of the total. The number of patents granted, 30,400, was 14 per cent more than in 1991.

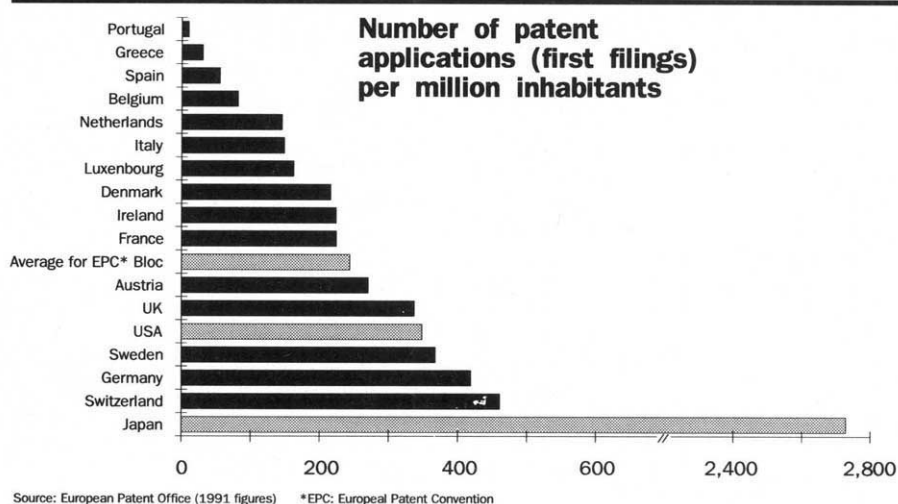
But Paul Braendli, president of EPO, warned that the cost of filing a successful patent in Europe is too high. At an average of DM48,000 (US\$38,000), European patents cost twice as much as US patents and four times as much as Japanese. Nearly half the cost — DM22,000 — is for translation into the eight languages of the European Communities.

The language rule should be changed, says Braendli, if small European businesses are to compete effectively. He believes that this may be one of the reasons why the average number of patent applications per head in 1991 in Europe was only 70 per cent

of that in the United States, and less than a tenth of that in Japan. He would have patents filed in their original language and translated only if challenged in court. This hap-

pens only, on average, for one per cent of patents. "Unnecessary" translating costs the patent office DM800 million a year.

Alison Abbott



EC hankers after more Framework research spending

Strasbourg. Increased spending on research and development in Europe won general approval at last week's summit meeting in Copenhagen, which marked the end of Denmark's six-month presidency of the European Communities (EC). This view, prompted by a paper on economic renewal in the EC by the European Commission's president, Jacques Delors, has now been echoed by the European Parliament.

A central feature of the Delors plan is that the EC should invest ECU5 billion (US\$6.1 billion) a year on a European information infrastructure. The document, commissioned by the member governments, says that an efficient decentralized economy requires a sound system of information highways and that the system should be continuously developed and updated, at an estimated cost of ECU5-ECU8 billion a year.

Member states have until 1 September to comment on the plan, so that a final document can be discussed by finance ministers under the chairmanship of Belgium, which now assumes the presidency of the EC. Decisions could then be taken at the next Community summit in December.

Danish prime minister Poul Nyrup Rasmussen, reporting to the European Parliament immediately after the summit, said that heads of state had responded positively to the Delors plan, which also urged member states to raise spending on research, development and innovation from 2 per cent of gross national product to 3 per cent.

Research was already on the agenda in Strasbourg, where the members of the European Parliament (MEPs) have been debating the plan to spend ECU18.4 billion on the 1994-98 Fourth Framework Programme. Pending ratification of the Maastricht Treaty (which gives the parliament more influence on EC research policy), members complained that the EC's commitment to research still falls short of the target of 6 per cent of their budget, agreed as long ago as 1985.

Indeed, Rolf Linkohi (Social Democrat, Germany), who is rapporteur of the Energy, Research and Technology Committee, claimed that the proportion of the EC budget to be spent in the new framework programme is a reduction from 4 to 3.9 per cent.

The parliament also renewed its demand that renewable forms of energy should have the same priority as thermonuclear research, with more research on the environmental impact of waste disposal. Last week, it also endorsed a draft directive that, within ten years, 90 per cent of the 50 million tonnes of packaging waste produced in the EC each year will be recycled or incinerated to produce energy, a proposal unlikely to be backed by the Council of Ministers. **Arthur Rogers**

Indian libraries pool science journal subscriptions

New Delhi. Universities and research laboratories in India, forced by lack of funds to cut out subscriptions to foreign journals, are now networking their science libraries.

Under this scheme, no two libraries will import the same periodical, and all journals will be pooled. A scientist who cannot find the journal he wants on his own campus can get a photocopy of any article he needs from another library in the network.

Researchers complain that shrinking libraries are slowing down their work. The number of foreign journal subscriptions at Indian science libraries has fallen from nearly 22,000 in 1983 to 9,500 as a result of budget cuts, increased prices and erosion of the value of the rupee.

For example, the Indian Institute of Science in Bangalore and the Tata Institute of Fundamental Research in Bombay have reduced their subscriptions by a third and the Indian Institute of Technology (IIT) in New Delhi has dropped half the foreign journals on its subscription list. The situation is even worse in universities and in the 40 laboratories of the Council of Scientific and Industrial Research, which says it has funds only to pay salaries.

"Resource sharing through computer interconnection is the only way to cope with the information drought in academic and research institutions", says Dr T. Viswanathan, director of the Indian National Scientific Documentation Centre (INSDOC), which is leading the network movement. The first network, which started last month in Madras, already connects ten libraries. They exchange the contents pages of periodicals, and scientists order the articles they want electronically. A courier service run by the network centre delivers a photocopy within 48 hours.

"By paying only a \$7,000 annual membership fee, a library has access to all the foreign journals collectively subscribed to by all the network members", says

Viswanathan. The libraries save about \$200,000 a year by not duplicating subscriptions. According to Viswanathan, all the 60 science libraries in and around Madras will have joined the network in the next few months. In Delhi, 28 libraries have formed a network and have saved \$170,000 by pooling journals. Similar networks are being set up in Bangalore and Calcutta.

Researchers whose libraries do not belong to a network are being encouraged to subscribe to INSDOC's Contents Abstracts and Photocopying Service (CAPS). For a nominal fee, subscribers receive every month the contents pages of up to 40 journals of their choice by fax or e-mail, and can then order abstracts or complete articles. INSDOC's own library can service 20 per cent of requests. "If we don't have the journal, we get it from a library that has", says Viswanathan.

INSDOC has regional offices in Calcutta, Bangalore and Madras, and a van makes a daily round of local libraries to pick up journals and photocopy the articles on the spot. "If the subscriber is in Delhi and the journal is also in Delhi, the photocopy will be hand-delivered within 48 hours", says Viswanathan. "Otherwise it will take up to a month." INSDOC has links with 800 science libraries and has access to every foreign science journal imported into India.

According to Viswanathan, most CAPS users are university and industrial scientists, and more than half the requests are for articles in medicine, the life sciences and engineering, most of them from *Nature*.

But most scientists in India believe that the problems caused by lack of access to journals can be only partly solved by networking and CAPS. "Networking of national science libraries will take time", says Mohinder Singh, chief librarian of IIT. "This cannot immediately make up for the steep fall in the acquisition of journals and books."

K. S. Jayaraman

Science network for Europe

Munich. Science information officers representing national science councils and major research institutes of ten European countries have organized a communication network to exchange information on science and science policy. Known as ESCIN (European Science Communication and Information Network), the group plans to publish a European handbook of scientific organizations this year, and is negotiating the start-up of a European-wide newsletter.

Communication in Europe is patchy. Countries tend to be unaware of major national scientific events and specific national laws affecting the practice of scientists in other countries. Founder member Hein Meijers, from the Netherlands Organization for Scientific Research (NWO), says that there has for a while been a strong feeling that scientists needed to communicate between themselves "at a non-political level". ESCIN is hoping for financial support from the European Communities.

Alison Abbott

Autonomy by stealth for French universities

Paris. Thirteen conservative deputies last week provoked a rerun of an old debate on the status of higher education in France; they are promoting a bill in the National Assembly that would give universities a modestly greater say in their own affairs.

This seemingly trivial private members' bill has reopened the traditional struggle between those who maintain that the homogeneity of the Napoleonic university system is an obstacle to progress (the Girondins) and those who see a plot to lead France down the road to an Anglo-Saxon model of competitive universities with restricted entry (the Jacobins).

Successive left-wing governments have attempted to reform the university system while keeping its nationalized status — the most recent with some success — while the right has regarded privatization as a remedy. Both have been impeded by the resistance to change of powerful lobbies of students and *professeurs*.

Francois Fillon, the new conservative minister for higher education and research, seems bent on continuing this political ping-pong, but by stealth. Alain Devaquet, Fillon's counterpart in the 1986 conservative government, resigned after making public plans to end the cherished principle of open access and to let universities set their own fees; there were riots, during which police killed one student accidentally.

So Fillon has emphasized the continuity of his policies with those of the former Socialist government, playing down talk of change. "I remain attached to pragmatism", he says. But doubts as to his intentions surfaced last week in the National Assembly, during debate of the contentious bill, which he supports.

The bill would let universities opt out of the 1984 Savary law, which specifies national norms for such things as the composition of a university council, course structures and basic curricula, the organization and status of associated laboratories and financial controls. It would also extend indefinitely the three-year dispensation already accorded to the seven new universities created over the past two years.

Opting out would allow universities to experiment with new forms of organization, perhaps cutting the number of classes for a subject or switching to a tutorial system. Fillon says this would allow universities to adapt themselves to local needs. But for the time being he discounts the possibility of universities opting out completely, saying that he will consider requests for "experiments" only on a case-by-case basis, and not if they were "contrary to the objectives of a university, or would jeopard-

ize its national character".

Fillon expects no more than a dozen opt-out requests this year. By choosing this system, which requires a majority vote by a university council, instead of imposing change, Fillon may ensure that reform gains momentum gradually, while avoiding dissent.

In last week's debate by deputies, Julian Dray (Socialist) described Fillon's strategy as "the art of camouflage". Why, if the bill was as harmless as the government claimed, had it chosen to debate it when the students were sitting examinations? Dray accused Fillon of "chasing the myth of America's competitive universities," and asked, "but for one Berkeley or Harvard, how many state universities are without talent?"

Officials at the higher education union (SGEN-CFDT) are also unconvinced by Fillon's promise to maintain financial equality between universities in different regions,

to uphold the national degree and to forbid universities from raising course fees. "We don't want to create a hierarchy, where a degree, from say Strasbourg, has more value than one from somewhere else," they say; "it's the thin end of a wedge towards creating an Anglo-Saxon system, which we do not want."

Ironically, Fillon's plans for university autonomy are sustained by a system introduced by Socialist minister Lionel Jospin, which predetermines a university's activities in a four-year contract negotiated with the ministry. That allows a university to spend its budget as it chooses and gives it the freedom to be different.

Fillon has further strengthened the principle by uniting the previously distinct contracts for research, teaching and general expenditure. The final piece in Fillon's plans for greater university autonomy is legislation, expected to be in place by autumn 1995, that would devolve day-to-day administration, including control of staff careers and budgets, from his ministry to each university. "It's ridiculous," he says, "that every decision, no matter how small, needs permission from Paris."

Declan Butler

ESF in quandary over EC relations

Munich. The search by the European Science Foundation (ESF) for a more relevant role in the new Europe continues, but last week's council meeting at Strasbourg came no nearer to deciding exactly what it should be.

ESF represents 59 national research councils and institutions in 21 countries, nine more than the membership of the European Communities (EC). At present, it spends its FF60 million annual budget on two major activities: science programmes and science (people) networks. It is not a grant-giving agency, but instead provides funds to facilitate communication and collaboration between research groups in different countries. The ESF's third major activity — 'Euroconferences' — is supported by a further ECU1.1 million grant under the EC Framework 3 programme (1990–94).

Relations with the EC are at the root of ESF's soul-searching. Last year, it embarked on a 'strategic reassessment' at the urging of French members, who argued that the 20-year-old foundation should match its goals to the changed political circumstances in Europe, and should in particular seek to complement EC research activities with more clearly defined strategies. Member states were asked to put forward their suggestions; it is still hoped that there will be agreement in time for the general assembly in November.

Last week's council meeting brought consensus only slightly nearer. Internal changes have been the more easily agreed. Thus there is general agreement to strengthen

the influence and decision-making power of the standing committees, which judge the quality and relevance of grant applications. The variable importance of these committees has long been a sore point.

On the other hand, complementary links with the EC are seen as a double-edged sword. There is agreement that the ESF should have a more formal advisory function at the European level, and that there should be links with the EC allowing ESF's independent advice to be efficiently sought and given.

But some members fear that the foundation could end up playing a role subservient to the EC, thereby allowing its fiercely guarded political independence to be eroded. The suggestion that ESF could provide a grant refereeing service for the EC and smaller member nations has been coolly received for the same reason, and also because individual countries believe they can manage with the refereeing arrangements they already have.

This introspection takes place against a general background of change at Strasbourg. Peter Fricker, former head of Switzerland's research council, the Nationalfonds, took over as secretary general in April for a five-year period after the retirement of the United Kingdom's Michael Posner. And the president, Umberto Colombo, is planning to step down in November because of his appointment in May as Italy's minister of universities and science. Nominations are being solicited for his replacement. **Alison Abbott**

CERN jumps another hurdle towards proton accelerator

Munich. Carlo Rubbia, director general of the European particle physics laboratory in Geneva (CERN), has confirmed that he will be ready by December with a detailed budget for the proposed next-generation particle accelerator, the Large Hadron Collider (LHC). The news will be a relief to the many who feared that technical difficulties in the development of the ultra-high performance superconducting magnets would delay the project.

Rubbia told last Friday's meeting of the CERN council that the magnets being developed under CERN contracts in seven different member countries would, in the next few months, be shown to be capable of achieving the required magnetic field of 9 Tesla. CERN hopes to use the LHC to search for high energy particles — the top quark, the Higgs particle (or 'God particle') — and supersymmetrical particles, which are predicted to exist, but which have not so far been detected.

But fixing a budget is, of course, only the very start of the story. The cost of the accelerator is estimated to be around US\$3 billion, and CERN's 17 member states will need to be convinced that the return will be worthwhile before they agree (as they must unanimously) to underwrite such a large-scale project. US congressmen were not convinced that a similar, but much more expensive proton collider (the Superconducting Super Collider or SSC), to be built in Texas, could be justified by its esoteric research aims. They voted last Thursday against financial support (see right).

The LHC, if it gets the go-ahead, could become even more international than now planned if the SSC is abandoned. CERN has an open-door policy to scientists in non-member countries; this year 500 of its 6,000 visiting scientists were American. "You can't marry a wife until the husband is dead", says a CERN spokesman, "but CERN could benefit from a more collaborative effort if only one proton collider is built."

The CERN council meeting also confirmed the election of Hubert Curien as CERN president in 1994. Curien was professor of physics at the University of Paris and l'Ecole Normale Supérieure before working for two spells as research minister in the French governments in 1984–86 and 1988–93. He takes over from Sir William Mitchell for an initial period of one year.

Christopher Llewellyn Smith, professor of theoretical physics at the University of Oxford, takes over from Carlo Rubbia as director general for a five-year stint at the beginning of next year. He is a firm supporter of the LHC programme.

Alison Abbott

SSC falls victim to congressional austerity

Washington. The Superconducting Super Collider (SSC) was probably mortally wounded this week, when a proposal to pay \$620 million towards the Texas accelerator in the coming year was defeated in the US House of Representatives. Supporters of the \$9-billion project in Congress and the high-energy physics community fear that, even if the Senate supports the SSC (as is probable), last Thursday's 280 to 141 majority will make it impossible to rescue the project when representatives of both houses of Congress meet to decide the matter finally.

The SSC vote was only one manifestation of the wider mood in Washington during a week that saw big science take a hammering from congressmen acting tough to cut budget spending:

■ On Tuesday, authorization of President Bill Clinton's revised \$10.5-billion space station plan scraped through the House by one solitary vote.

■ Two days later, as the SSC fell, the House voted 272 to 162 to cut to zero the funds requested for the Integrated Fast Reactor (IFR) at the Argonne West laboratory in Idaho. The proposal before the House already reduced the IFR's budget from \$125 million to \$32 million (see *Nature* **362**, 383; 1993).

■ By a majority of 329 to 91, the House removed a \$25 million request to continue development of the SP-100, a small space-based nuclear reactor left over from the Star Wars programme.

Supporters of the SSC complain that there was little they could do, in the prevailing atmosphere, to get their arguments heard. "They want to be able to go home and say 'we've cut something'", one official said.

According to congressional staff, the collider has suffered because many in Congress see their choice as being between the SSC and the Space Station. The latter is said to support 75,000 jobs, many of them in the battered aerospace industries of Texas and California. SSC is thus easier to cut, although cancellation will mean \$1.5 billion wasted and a spectacular white elephant near Dallas.

Leading high-energy physicists who have fought for the SSC for almost a decade are almost ready to throw in the towel. "It seems to me that we have lost," concedes Stan Wojcicki of Stanford University, California, chairman of the Department of Energy's High Energy Physics Advisory Panel (HEPAP). "I worry about what it says about our nation, and its ability to carry things through", Wojcicki says.

He concedes, though, that scientists will have themselves to blame if the project falls. "A lot of us don't talk enough about what we

are doing, so the public and its representatives are not aware of it." Wojcicki, now on sabbatical at the SSC laboratory near Dallas, says that, despite his own pessimism, most physicists are continuing on the basis that the SSC will be built. "You can't base your actions on just one vote," he says.

Supporters of the collider have also attacked enemies in Congress for stretching out the programme, thus increasing its cost and scaring off international collaborators — and then attacking it for having higher costs and no collaborators.

In a similar vein, in the wake of the Space Station's narrow victory, George Brown (Democrat, California), who is chairman of the House Science, Space and Technology Committee, chastised his opponents for "delaying, obstructing, and driving up costs on this project". Brown — doubtless to no avail — asks Space Station opponents to hold fire on a programme that, "has been debated, amended and redesigned nearly to death."

The SSC debate will now rest until late August when the Senate, which last year supported it by 66 to 34, will probably back it by a narrower margin. The conference between both houses — which saved the programme last year, in more auspicious circumstances — will then decide its fate.

The Space Station faces another hurdle this week with a second House vote on appropriation of funds for next year. With strong Senate and White House support and NASA engaged on finalizing the president's preferred design option, its prospects look more secure than they have done all year.

But a planned visit to Washington by Russian prime minister Viktor Chernomyrdin to discuss fuller Russian participation has been postponed, because of US concerns about Russia's supply of missiles to India.

In Idaho, the fast breeder reactor could yet survive the House vote against it. The Argonne West facility has strong backing in the Senate, which may even support full development funds as opposed to the 80 per cent cutback in the president's budget. A compromise might then follow.

But the space reactor, not included in the president's budget but added later by a House appropriations subcommittee, is well and truly dead after last Thursday's vote. Taken as a whole, these votes reflect more than just a fashion for deficit-reduction. They may prove instead to be an early indicator of how vulnerable US science funding has become without Cold War arguments to protect it. If that is the case, researchers engaged on projects that are not 54 miles incircumference may soon feel the consequences.

Colin Macilwain

Stanford and Cornell in the frame for B-meson factory

Washington. An expert panel will this week assess whether a high-energy physics facility known as the B-meson factory should be built at Stanford (California), at Cornell University (New York) — or not at all.

The 12-strong panel, chaired by Dr Stanley Kowalski of the Massachusetts Institute of Technology, visited the Stanford Linear Accelerator Center (SLAC) last week and then went straight on to Cornell, quizzing rival project teams on their proposals. The panel will report to the administration next week; Congress has asked for a choice of the best option by 15 July.

Some members of the high-energy physics community are worried that these tight deadlines will not be met, causing the Congress to withdraw the \$36 million it has been planning to allocate to the project in the coming year. But even as the House of Representatives last week moved to axe funding for other programmes (see opposite), funds for the B-meson factory project survived the appropriation process intact; Congress makes a final decision later in the summer, if and when a site has been chosen.

The purpose of the proposed facility is to create large numbers of B mesons, which can be produced by colliding electrons and positrons. Existing colliders produce B-mesons, but as the particles of opposite charge usually have identical energy, the mesons emerge in random directions. The new objective is to produce a beam of B-mesons, for which purpose unequal energies are necessary. The B-meson factory will therefore tuck a second storage ring, carrying positrons of lesser energy, next to an existing ring at either SLAC or Cornell.

Interest in the production of B-mesons centres on the likelihood that their decay will throw light on the violation of parity and charge-conjugation invariance in the weak nuclear interaction — phenomena which may be crucial in explaining the predominance of matter over anti-matter in the known universe. Because of their large mass, the B-mesons will also be prolific sources of other nuclear particles (hadrons).

Stanford's play for the project is due to lobbying by SLAC's director, Burton Richter, of Californian congressmen, who won funds for it in President Bill Clinton's first budget. SLAC says it would update its 800-metre diameter Positron-Electron Project (PEP) ring (now 13 years old) at a cost of \$161 million over four years. B-meson detectors would cost a further \$59 million — half from overseas collaborators.

SLAC, a DoE laboratory with 1,400 staff and an annual budget of around \$130 million, needs the B-meson factory to keep its researchers busy during a slack period be-

fore it can renew its main linear accelerator.

The rival bid is cheaper. Cornell says it could upgrade the Cornell Electron Storage Ring (CESR or 'Caesar') to a proper B-meson factory for just \$110 million, although but SLAC says the two sets of figures are not directly comparable.

Kowalski insists that the administration and the Congress, not his panel, will eventually choose the winner. He says the panel's job is to evaluate the two proposals "to see if they make for good physics" and to say if the cost estimates "are credible and complete".

NSF and DoE will have a week jointly to decide the outcome, but the Office of Management and Budget and the Office of Science and Technology Policy as well as various congressional interests are certain to take an interest. Just who will make the choice, and how, is unclear: "That's the million dollar question", says Kowalski. "I don't believe there is a formula."

But some observers believe there is a formula: put the B-meson factory in California, where DoE wanted it in the first place. The decision will be "political", according to a prominent physicist who declined to sit on Kowalski's panel.

Karl Berkelman, director of CESR, says on the other hand that he will not be "greatly surprised" if the project is put back on hold. But some supporters are more optimistic: the B-meson factory has few enemies, its costs are manageable and its objectives clear. And if the Superconducting Super Collider is indeed to perish, this more self-effacing project may benefit from Congress's subsequent pangs of guilt.

Colin Macilwain

Push for TB vaccine

London. The pharmaceutical company Glaxo is to provide £10 million for a five-year university-based programme of research into tuberculosis — the world's main killer infection, which claims three million victims every year — to be carried out in Britain and South Africa. The funds will support basic research on the development of new vaccines, using the techniques of molecular biology and genetic engineering, and of new antibiotics effective against strains of the tuberculosis bacterium that are resistant to existing treatments. Research institutions collaborating with Glaxo in Britain will be the London School of Hygiene and Tropical Medicine, St Mary's Hospital Medical School and the University of Birmingham. The South African research groups will be based at the University of Cape Town, the University of Stellenbosch and the South African Medical Research Council. **D.D.**

Yeltsin's two steps to Eureka club membership

London. Russia's application to join Europe's advanced technology research programme Eureka was approved in principle last week, but several conditions must be met by the Russian government before membership can become a reality.

Eureka was set up at the initiative of the French government in 1985, and now covers more than 800 separate projects involving industrial companies, public laboratories and university research groups in 20 European states. Russia's application was formally agreed by research ministers from the member states in Paris. Among the potential benefits of membership is that of speeding the conversion of Russian military research laboratories to civil goals.

One condition to be satisfied by Russia is that its laws on patents and intellectual property, previously based on the idea that the rights to scientific discovery belong to the state, should come into line with those in Western Europe; President Boris Yeltsin has undertaken to submit a decree to this effect. He is also required to set up procedures to ensure that civilian technologies imported from the West are not diverted to military purposes, which may be more difficult.

Ironically, Eureka was inspired by the prospect that US industry would benefit from civilian spin-offs from President Reagan's proposed Star Wars programme.

Despite initial scepticism, Eureka is now flourishing. At last week's meeting, 193 new projects, each of which must involve research groups from at least two participating states, were announced, bringing the total investment in Eureka projects over the past eight years to £12 billion.

An evaluation of Eureka's social and economic impact by experts from 14 countries and co-ordinated by the University of Manchester's programme of Policy Research in Engineering Science and Technology (PREST) shows over 40 per cent of participants in individual projects expect a substantial increase in sales over the next three years, and that 88 per cent will introduce new products. The study confirmed that cooperation tends to be "vertical" (between suppliers and customers) rather than "horizontal" (between competitors).

According to Luke Georgiou of PREST, the evaluation also showed that public funding had been an essential element in ensuring the success of Eureka, in particular by motivating companies to join individual projects. Russia's application for membership follows Hungary's acceptance as a full member last year. In addition, national information points have been set up covering 12 countries in Central and Eastern Europe.

David Dickson

British research council jumps at gene therapy

London. The British Medical Research Council (MRC) is pressing ahead with a commitment to stimulate gene therapy in Britain, and has named a 20-member committee of academic and industrial scientists and government officials to promote research and its application in the field.

This development follows shortly after the decision by an advisory committee to approve a second gene therapy trial, involving cystic fibrosis patients. But it coincides disappointingly with an acknowledgement by those carrying out the first trial that their efforts have so far failed to produce significant improvement in the condition of a one-year-old girl with adenosine deaminase (ADA) deficiency.

The chairman of the MRC committee will be Peter Rigby, head of the genes and cellular controls group at the MRC's National Institute for Medical Research. Its official terms of reference are "to give advice, resolve issues and problems, co-ordinate clinical trials and the resultant data, and encourage interaction with industry".

Two-thirds of the committee's members are scientists working in universities, charity-funded laboratories and research council institutes. Others include representatives of the drug companies Glaxo and Zeneca, and officials from the Medicines Control Agency and the Department of Health. The committee does not include independent experts in medical ethics, reflecting MRC's belief that responsibility for regulating gene therapy should rest with the Department of Health.

One of the new committee's first tasks, according to Diane McLaren, executive secretary of the MRC's molecular and cellular medicine board, will be to allocate funds for a number of university-based gene therapy centres under a government initiative on genetic approaches to human health; 27 outline proposals have been received since applications were invited earlier this year.

In contrast with the speed of MRC's move and the enthusiasm of research groups, the British government continues to drag its heels over regulatory procedures. Although it is 18 months since a committee chaired by the lawyer Sir Cecil Clothier recommended a special gene therapy advisory committee (see *Nature* 355, 190; 1992), the full committee is not now expected to be announced until September. But it is known that it will be chaired by Dame June Lloyd, who retired last October as Nuffield professor of child health at the University of London's Institute of Child Health, at which the ADA experiments were planned. Lloyd, a specialist in familial lipoprotein disorders, is scientific adviser to the Association of Medical Research Charities, and a member of the

MRC committee.

Until the new body comes into operation, proposals for gene therapy experiments are still being assessed by the Clothier committee. At its last meeting, the committee approved a set of clinical trials planned at the Royal Brompton Hospital and St Mary's Hospital Medical School into a potential treatment for cystic fibrosis.

The final go-ahead for the experiments, which will use liposomes to transfer functioning DNA into volunteer patients through nasal inhalation, awaits detailed approval by the Medicines Control Agency. But Bob Williamson, professor of biochemistry at St Mary's, says the trials are expected to start "within the next few months".

A second group, based in Oxford and

Cambridge, which described earlier this year the insertion of a cystic fibrosis gene into mice bred with a defect in the gene, is also hoping to start trials on human patients before the end of the year.

Several experiments in the gene therapy of cystic fibrosis are in progress in the United States, using adenovirus vectors, but there are reports that some of these have run into difficulties related to viral pathogenicity. The British researchers hope the use of liposomes will open up a more benign route to successful treatment, answering scepticism about the viability of liposomes as an alternative to adenoviruses by suggesting that the optimal therapy may eventually turn out to be a combination of approaches.

Meanwhile, physicians at the Great Ormond Street Children's Hospital say that, despite the lack of evidence that the ADA gene replacement therapy carried out on one-year-old Carly Todd has been successful, there is still 'a possibility' that treated stem cells will develop into a new immune system for her.

David Dickson

Max Planck expands toward the east

Munich. The Max Planck Society (MPG) has been tipped off by the federal government that funding for five new Max Planck institutes in the former East Germany may indeed be forthcoming, the general squeeze on resources precipitated by reunification notwithstanding. Last week's meeting of the MPG senate at Trier was therefore able to endorse its plans, agreed in principle last year, to set up in the new *Länder* three new institutes in the natural sciences and two in the social sciences.

Uncertainties over finance have delayed the setting up of the institutes. The federal funds provided in the 1990 reunification treaty for building up research in the new *Länder* did not stretch as far as funding all seven institutes originally earmarked. But setting up institutes virtually from scratch has also required more time than had been predicted. Only two — the Institute for Microstructure Physics at Halle and the Institute for Colloid and Surface Science in Berlin — are so far established.

There are now signs that the federal government is willing to provide its share (half) of the costs of the five new institutes. Research minister Paul Krüger, himself an east German, is firmly committed to giving priority to initiatives in the new *Länder*, and is this week negotiating with finance minister Theo Waigel in preparation for the federal budget on 13 July.

The other half of the budgets of Max Planck institutes normally comes from all the German *Länder*, but the reunification treaty requires only the new *Länder* to contribute to costs for their own institutes, at least until 1995. Thus the new institutes will cost western German *Länder* nothing for the time being, giving the lie to complaints that

the east is "draining resources" from the west. Most of the new *Länder* are said to be ready to sign contracts.

The new institutes include:

■ An institute for **infection biology** in east Berlin, to be established by the end of 1994. It will eventually be housed in a new building near the Charité university clinic to be built by 1998. It will have 100 positions, 35 for scientists, and an annual operating budget of around DM11 million (US\$7 million).

■ An institute for **molecular plant physiology** near Potsdam in Brandenburg. Work will begin in 1994, but the 90 staff (26 scientists) will move into a new building in 1998. The institute will have a budget of about DM9 million a year.

■ An institute for the **physics of complex systems** at Dresden opens next month, and moves to a new building on the campus of the Technical University in 1998. Although small (37 positions, of which 17 will be scientists), it will be one of the few in the world dedicated to this multidisciplinary field; an extensive guest programme is planned.

■ Institutes for **economics** (Jena) and **history of science** (Berlin) will increase the number of Max Planck institutes in the social sciences, which have been much neglected. The new institutes are much welcomed in the east, where unemployment among scientists has been high — and confidence low — since reunification. However, all three directors so far designated are from western Germany, matching current trends (see *Nature* 362, 685; 1993).

But the good news has not ended. MPG says that it is considering further initiatives in the new *Länder* for the second half of the 1990s.

Alison Abbott

A future for South Africa's nuclear agency?

Pelindaba. When South Africa's new government takes office in May next year, one of its first tasks will be to assess the future of the three strategic energy industries created by the National Party government. Two of these, SASOL and MOSSGAS, recover oil from coal and natural gas respectively. The third is the Atomic Energy Corporation (AEC), which occupies a vast site at Pelindaba, 20 km west of Pretoria in the foothills of the Magaliesberg mountains.

This year, the AEC will receive a R469 million (US\$142 million) government subsidy, a third of which comprises interest and loan repayments guaranteed by the state. A new government can but honour these repayments if it wishes to remain creditworthy. The question is whether it will continue to subsidize operating costs to maintain a nuclear industry, or cut its losses by simply writing off the capital repayments.

The AEC's own programme aims to make the organization 75 per cent self-sufficient, in terms of running costs, by the turn of the century. Exempt from tax, last year it generated just under 30 per cent of its operating budget, or R177 million, from sales, just under 60 per cent of which were of nuclear fuel products. The balance comprised irradiation products and services, fluorochemicals and industrial products arising from the corporation's expertise in aluminium welding. But now, the AEC sees its salvation in the non-nuclear sector: it hopes to double its sales in the current financial year.

The highly enriched uranium (HEU) used to manufacture South Africa's nuclear devices, now back in the hands of the AEC, is being used to re-load the SAFARI-1 reactor. Commissioned in 1965, that was designed to run on 93 per cent HEU, but has been running on 45 per cent HEU since the United States stopped supplying SAFARI in 1976. The bomb material will be sufficient to supply the reactor for the rest of its lifetime, estimated at between 15 and 20 years.

The two main uses of the reactor are the production of medical and industrial isotopes, and the irradiation of silicon for the manufacture of power-rectifiers. The reactor is currently operating at about one-fifteenth of its capacity in each area, and last year produced sales of only R5 million from isotopes and R1 million from silicon irradiation. The problem is that the South African market for isotopes is small and that for silicon irradiation almost non-existent. But Don Mingay, SAFARI's manager, hopes to be able to increase exports by capitalizing on a predicted global shortage of neutrons when the Belgian BR2 reactor closes in 1996.

The nuclear fuel section contains three plants: one for conversion of feed material

to uranium hexafluoride, the Z-plant, where enrichment is performed by the stationary-wall centrifuge process (developed in South Africa), and a fabrication plant, where fuel rods are manufactured.

The Z-plant appears doomed: the chief of the AEC's nuclear fuels division, Piet Venter, concedes that it is too energy-intensive to be viable under any circumstances. Even under an agreement this year by which the Electricity Supply Commission (ESCOM) supplies the AEC at a cheap rate, the plant is expected to generate income



Dr Anthony Jackson, head of the AEC's business development unit.

from sales of only R80 million, compared with operating costs (excluding capital depreciation) of R134 million.

Nobody will say exactly how much energy Pelindaba's Z-plant consumes. But it is a remarkable irony that the AEC uses one per cent of ESCOM's total electricity production, only 3.3 per cent of which is generated by Koeberg, its single nuclear power station. Venter says that the technology would have been economically viable if the plant had been constructed to produce five million separative work units (SWU), as originally planned. Currently, it is producing 300,000 SWU annually, which represents three-quarters of its full capacity.

Somewhere along the line, a decision was taken to scale down the capacity dramatically, but it appears that its cost-effectiveness was never reassessed. More likely, this was not really the issue: the project was essentially an expensive smokescreen for

the pilot plant, which used the same process (on a smaller scale) to enrich uranium for weapons manufacture (see *Nature* 362, 384; 1993).

The decision on whether to close the Z-plant may even end up on the agenda before the government changes hands next year. The fate of the conversion and fabrication plants is less clear-cut. Both would be viable if South Africa were to build a second nuclear power station, but that is unlikely before 2000, even if the new government deems it politically acceptable.

Although both plants could break even through exports, these are prohibited in the case of the fabrication plant; it is a condition of the technology transfer that fuel elements can be sold only in South Africa. The AEC has applied for a review of this provision, which is currently being considered.

Another potential area for cutback is in the AEC's R80-million research and development budget, half of which is at present being spent on the development of a molecular laser uranium enrichment process. Current plans are to fund the project for two more years at this level, with a prototype due to be running by 1995. The corporation had hoped eventually to replace the Z-plant with one using molecular laser enrichment.

But AEC's chief executive, Waldo Stumpf, says that the future of the project will depend on whether the private sector will finance it on a commercial basis. He claims that the AEC has overcome three major technical problems in molecular laser enrichment: the micro-aerodynamics of flow cooling, repeating laser pulses at 2 kHz and controlling uranium hexafluoride concentration in the flow-cooling network. But others, both within and outside the AEC, ask whether another expensive artefact of South African isolation is being nurtured.

Can the AEC turn itself into a viable organization? Both the corporation's management and its board, says Stumpf, are "very painfully aware of the non-commercial decisions and over-investment in the middle seventies". The AEC has cut its staff from more than 8,000 in 1986 to a little over 3,000 today. But opulence persists: the Pelindaba parking-lot is replete with luxury models of German motor cars.

Michael Cherry

Correction: ESO negotiates with Chile over agreement on telescopes

This news article (*Nature* 363, 384; 1993) incorrectly reported that the European Southern Observatory (ESO) site at Paranal in Chile was bought by ESO twice, from different private owners. This is true of the La Silla site; the Paranal site was donated to ESO by the Pinochet government. □

ACOST never browbeaten

SIR — Your leading article (*Nature* 363, 381; 1993) on the White Paper on British research policy casts doubt on the independence of the Advisory Council on Science and Technology (ACOST) and its successor, the Council for Science and Technology.

During my three-year term of office as chairman of ACOST, I can say categorically that the council has never been "browbeaten by determined ministers and their chief scientific advisers". ACOST has addressed the issues it wished to address and has said what it wanted to say. Certainly, in selecting issues to study, ACOST has taken note of government as well as public concerns, and in suggesting mechanisms for implementing recommendations, ACOST has taken note of government policies. To do otherwise would be foolish and would not be more "independent".

I welcome the establishment of the Council for Science and Technology as a positive development of the role that ACOST has undertaken. The fact that it is to be chaired by the Chancellor of the Duchy of Lancaster, on behalf of the Prime Minister, will place advice on science and technology policy at the heart of the political process in government, which is where it belongs.

Within the Council for Science and Technology, I have no doubt that the independent industrial and academic members of the new council will be more than capable of presenting rigorous independent advice to ministers. Additionally, paragraph 2.41 of the White Paper makes clear the willingness of government to receive high quality independent advice from other sources outside government, not least the quaintly named but, hopefully, very effective National Academies Policy Advisory Group.

Robin Nicholson

Cabinet Office,
Advisory Council on
Science and Technology,
Albany House,
84–86 Petty France,
London SW1H 9ST, UK

Children of choice

SIR — According to James Watson¹: "The world must shed the idea that this is evil [abortion of fetuses diagnosed with genetic defects], as it is a true act of moral cowardice to allow children to be born with known genetic defects." The idea is an old one. In ancient Sparta, Olympic athletes were pressed into service to provide the fruit of their loins for the better-

ment of Sparta. Weak or malformed infants were exposed². That was no evil because it was done for the benefit of society.

Hermann Muller³, who won the Nobel prize for Physiology or Medicine in 1948, considered the benefits of this idea to society in detail. One of his objectives was to free women from the menial servitude of child-bearing and child-rearing by separating sexual intercourse from reproduction. Laboratory methods for culturing and transplanting human ova would "greatly extend the reproductive potencies of females possessing characters particularly excellent, without thereby necessarily interfering with their personal lives." Muller was the first to develop the technique of preserving sperm, human and cattle, in liquid nitrogen.

To be sure, Muller was aware that popular pressure might result in "a maximum number of Billy Sundays (a fundamentalist preacher), Valentinos (an actor), Jack Dempseys (heavyweight boxing champion), Babe Ruths (baseball) and even Al Capones (a gangster)". The contemporary roster of popular heroes is no less disconcerting.

Muller added that the selection of genetic material would be voluntary and that... "Compulsion need enter, if at all, only in a negative role, to prevent exploitation of the enhanced possibilities."

Muller's extreme naïveté jumps out at one. The meek may be blessed and inherit the Earth, but they won't keep it long. It is exactly those individuals who are "unduly egotistic, aggressive or paranoid" who will attain the power to choose the genetic heritage of "children of choice".

Hubert P. Yockey

1507 Balmoral Drive,
Bel Air, Maryland 21014, USA

1. Short, N. *Nature* 362, 783–784 (1993).
2. Haeckel, Ernst *The Wonders of Life* (Harper, New York & London, (1905).
3. Muller, Herman J. *Out of the Night* (Vanguard Press, New York, 1935).
4. Huxley, Aldous, *Brave New World & Brave New World Revisited* (Harper & Row, New York & London, 1932, 1946, 1958, 1965).

SI units defined

SIR — The letter¹ "SI units explained" itself demonstrates confusion about the meaning of SI units. Regrettably, the debate continues with little or no reference to the primary recommendations from authoritative intergovernmental organizations^{2–4}.

SI units are defined as including base units (kg, m, s, . . . , mol) and derived units (for example, kg s⁻¹, mol kg⁻¹, mol m⁻³). The International System of Units is the coherent system encompassing these units. The authorities also recognize decimal multiples of SI units and give rules on how they should be expressed (for

example, mg s⁻¹, μmol kg⁻¹, nmol m⁻³); these units are not actually classed as SI units, so that their use can be viewed either as a "retreat from SI units" or as a supplementary set of units. There are also "non-SI units which may be used with the SI units and their multiples"⁴ or "units outside the International System" recognized for "use with the International System" such as day (d), litre (L), tonne (t). Such units are legally recognized units in the European Communities and in the United States. Choice between them is a matter partly of following the rules in the international agreements but also of taking into account logic, common sense and convenience. For instance, it is more convenient to use an expression with one prefix than two: g m⁻³ or mg L⁻¹ rather than mg dL⁻¹.

The choice between "mg/dL" and "mmol/L" for cholesterol measurements, however, is a matter of choosing mass concentration and substance concentration of cholesterol; it is not a question of validity of units. That choice ought to be governed by the relative informativeness of the two kinds-of-quantity. In the words adopted by the 30th World Health Assembly (Resolution 30.39 of May 1977): "For a proper understanding of chemical reactions . . . the use of the mole is essential. The use of mass units . . . serves no purpose other than the purely arbitrary one of deciding whether or not a given value is greater or less than a certain reference value. The expression of concentrations . . . in body fluids in molecular terms also serves this purpose, but in addition gives valuable insight into the balance of the constituents. Such insight cannot be obtained from mass units." This comment was part of a resolution entitled "Use of SI units in medicine"; it can be viewed as a restriction, but it is better to see it positively as a sensible preference.

We urge authors and editors to make the effort to acquire a sound grasp of the definitions and utility of SI units. Although there may be a certain charm⁵ in a multiplicity of measurements, the advantages of consistency seem overwhelming.

J. C. Rigg

Pudoc, PO Box 4,
6700 AA, Wageningen,
Netherlands

S. S. Brown

West Midlands Regional Laboratory
for Toxicology,
Birmingham B18 7QH, UK

1. van Assendelft, O. W. *Nature* 362, 388 (1993).
2. *Le Système International d'Unités (SI) The International System of Units (SI)* 6th edn (Bureau International des Poids et Mesures, Sevres, 1991). [ISBN 92-822-2112-1].
3. *Quantities and Units. Pt 0. General principles ISO 31-0* International Organization for Standardization 1992 [Obtainable from national standardization institutions.]
4. *SI Units and Recommendations for the Use of their Multiples and of Certain Other Units ISO/DIS 1000* (International Organization for Standardization, 1990) [Obtainable from national standardization institutions.]
5. Asten, E. L. *Nature* 345, 506 (1990).

Russia's summer bout of optimism

To have survived this far is a triumph, perhaps also an augury for the future, but Russian researchers remain anxious about the long-term future.

St Petersburg. "The trouble with Russia is that it's governed by optimists", Dr Sergei A. Korolenko, from the Institute of Cytology of the Russian Academy of Sciences, said last week after a brief conversation about the economy. Then, by way of afterthought, "That's only natural. It's natural selection: all the pessimists have left."

Whatever the cause, pessimism is less palpable now than for three years, since Gorbachev's version of *perestroika* began to sour. For laboratories and whole institutes to have survived these years, let alone Russia itself, is now recognized as a triumph, even a sign that the undefined problems ahead may also be surmountable.

The optimism might, of course, have been moderated if the Korolenkos of this world had known that, on the very day they were exchanging pleasantries with a foreigner, their government was trying to mortgage its strategic stocks of heavy metals to pay a debt to a Swiss bank, or if they had seen the women and children from some southern republic, perhaps thirty altogether, huddled together as if thrown in a heap, alongside one of the station platforms from which the sleeper-trains to Moscow leave; the men were circling around, begging.

So optimism must be tempered with realism, the essence of which is that the hardships and humiliations of mere living are no less than in Brezhnev's time, although different. Inflation is the bugbear, the general level of prices having increased more than a hundredfold, but unevenly. The price of a metro-journey in Moscow has gone from 5 kopecs to 5 rubles since 1987, but minor luxuries (such as air travel or international telephone calls) are now pegged to the US dollar and are more than 1,000 times what they were. (In 1987, one dollar would buy 60 kopecs; last week, it would buy 2,000 times as many.)

Salaries have risen, but not as quickly. At the Institute of Organic Chemistry in Moscow, the average salary is about 20,000 rubles a month, just a little less than the 22,000 rubles offered last week to those who clean metro stations. But a dozen researchers earn extra by working as typesetters for the desktop publication of three chemistry journals, to be produced more or less simultaneously in Russian and English.

Inevitably, hardship is forcing people out of science, or out of Russian science. One physicist in St Petersburg said he had almost wept when one of his brightest post-diploma students (on the way to the equivalent of a PhD) had left to work in a hotel;

"He's now opening the door for guests, I think." Fixing things for foreign businesses is a favourite alternative to science, but there is a small army of research-qualified general traders and a sprinkling of people who have set up technical enterprises in Russia. Some have already become bankrupt.

The flight overseas preoccupies the research managers, but its pattern varies. At the Ioffe Physical Technical Institute in St Petersburg, the largest in the Russian academy system (now that the Lebedev Institute in Moscow has been split), only five people are known to have found permanent jobs overseas. And although a hundred people left last year, some probably for short spells, all the vacancies have been filled with younger people already on the institute's books, keeping the roster of active researchers at about 1,200. The Ioffe institute attributes that success to its close involvement in the teaching programme of the St Petersburg Technical University. But the managers say they will not be able to repeat that trick for more than a year or two.

"Active" may not be the right word. With the general squeeze on budgets and the need to keep as many researchers as possible above the bread-line, spending on equipment and supplies has declined drastically. Overseas visits have been curtailed by Aeroflot's new realism on air-fares.

On the other side of that coin is the sad case of the Euler Institute in St Petersburg, named after the mathematician who helped Peter the Great to found the Russian academy, and opened less than a year ago by Sir Michael Atiyah, president of the Royal Society of London. One day last week, it housed only its director, the distinguished theoretical physicist Professor Ludwig Faadeev, his administrator and their two secretaries. The plan to use the handsomely restored bourgeois palace for seminars lasting several months is in abeyance, awaiting funds with which to bring up to a dozen mathematicians from overseas. Faadeev plans to apply to the new Soros fund. Or UNESCO may be able to help a little (as it has done with the restoration of the palace).

Faadeev may be disappointed that the germination of his institute has been postponed, but he is "bitter" to have been deserted by most of his senior colleagues at the Institute of Applied Mathematics, where he is also director. As recently as last August, he insisted that his hand-picked people, although perpetually on the wing, would not abandon St Petersburg. But now they have gone, mostly to academic posts in the

United States. Faadeev says that when he meets them, they portray themselves as the diaspora of his institute. For his part, he does not have the stomach for beginning all over again, finding bright students and bringing them on. And there is the generation gap.

The Russian Academy can no longer solve all problems. Its direct support of institutes ranges downwards from 90 per cent (at the Ioffe institute) to as little as 25 per cent at the Shemiakin institute (active in biotechnology) in Moscow. The physics institutes are the more dependent, and complain that Russian industry (an alternative source of funds) is more concerned to keep its plants at work than to invest in research. The Russian Foundation for Fundamental Research has been a lifeline for many research groups; high hopes are pinned on the Soros fund, due to publish the conditions of its research awards (up to \$100,000 each) this week.

A curious ambivalence towards the academy is emerging. With institutes having to survive by the wits of their members, central control is less obtrusive than formerly, and so is less widely resented. On the other hand, the trappings of privilege (limousines with drivers, the vulgar and expensive new headquarters building in Moscow, \$400-a-night suites in Washington hotels), although attenuated, are at odds with the delay in paying salaries, now one or two months in arrears at many institutes.

Some of these grumbles will be aired at a meeting organized for this Wednesday at the Institute of Foreign Relations, originally called to complain that President Boris Yeltsin has abandoned his pre-referendum promise that scientists' and teachers' salaries would be maintained, and has now decreed a freeze of public budgets.

The most serious danger for the academy and thus for its researchers is that it will fall out of step with the government. Its response to the removal of Andrei Gonchar, the deputy president of the academy and its most effective member, from his interim job as president of the fundamental research foundation (see *Nature* **363**, 662; 1993) was to demand the dismissal of the able science minister, Boris Saltykov. For his part, Gonchar said last week that the government was "right" (in the sense of being within its rights) to act as it did. In an earlier conversation, Saltykov had remarked that the academy "has done nothing to assist the reform process" in the past few years, but has "looked after itself". There are the makings of trouble.

John Maddox

Cooperation wins and stays

Manfred Milinski

THE wrangling over plans to cut global emissions of greenhouse gases is but one instance of our inability to secure the common good because of short-sighted national interests. Countries, prisoners and animals have in common the constraint that no authority guarantees their private contracts of cooperation. They are caught in a 'Prisoner's Dilemma', the metaphor for the problem of achieving cooperation among egoists. During the past decade, however, evolutionary theorists¹⁻⁴ have found it more and more likely that egoists can win only by using cooperative strategies in the Prisoner's Dilemma. Now, on page 56 of this issue⁵, Nowak and Sigmund present the winner of computer simulations of extensive evolution: it is indeed a cooperative strategy, one which they call 'win-stay, lose-shift'.

In the Prisoner's Dilemma (see figure) the players can either cooperate or defect (not cooperate). If they cooperate, both do better than if they had defected. But if one player defects while the other cooperates, the defector gets a very high reward and the cooperator gets nothing. Each should defect no matter what his opponent does if they meet only once. So both end up with a much lower reward than they would have gained if they had decided to cooperate.

Nevertheless, some game theorists guessed that if the Prisoner's Dilemma is played repeatedly by the same players a strategy that is not only cooperative itself but can also elicit cooperation from the opponent could be possible. The problem was too complex for the winning strategy, be it good or bad, to be determined with algebra. Axelrod^{1,2} had the idea of using evolution as a tool. If natural selection is efficient enough to produce the optimal design we observe everywhere in nature, it should have provided animals with the best strategy for the iterated Prisoner's Dilemma. He simulated evolution in computer tournaments for which he solicited strategies from game theorists. The strategies were paired against each other and they 'reproduced' according to their gains. After 1,000 generations a clear winner stood out: surprisingly it was not one of the perfidiously defecting strategies but a very simple cooperative one called tit-for-tat (TFT). This strategy starts cooperatively and then repeats the opponent's previous move. Thus two TFT players always cooperate but TFT is not

the sucker on encountering 'Always Defect'. Axelrod did not include mutant strategies coming up by chance, which is a prerequisite for natural evolution. Instead, he relied on the intuition of experts. (Perhaps a completely different strategy would have won if Einstein had entered.)

The glorious superiority of TFT wanes, however, when the players can make mistakes: one occasional misunderstanding



		Player 2	
		Cooperate	Defect
Player 1	Cooperate	$R=3$	$S=0$
	Defect	$T=5$	$P=1$

The Prisoner's Dilemma is everywhere. Two players (for instance riders on a tandem) can each decide whether to cooperate or defect. The matrix shows that player 1 gains more by defection not only when player 2 cooperates (5 instead of 3 points) but also when player 2 defects (1 instead of 0 points). Because this is true for either player, both will defect and end up with 1 point each although they could have gained 3 points each if they had cooperated. Hence the dilemma (and the reason why tandems have never become very fashionable).

ing condemns two TFT players to mutual backbiting until the next mistake occurs. This weakness of TFT is due to its deterministic nature, so Nowak and Sigmund³ carried out new computer simulations with a mixture of strategies each with a different set of probabilities (for example to cooperate on the first move with probability 0.7, with 0.9 after the opponent's cooperation and with 0.2 after his defection). The result was frustrating. Always Defect won in nearly all simulations. But things changed dramatically if one of the initial strategies (added deliberately or by chance) was close to strict TFT. Then a new champion emerged — 'Generous TFT'. Like TFT this strategy starts with cooperation and responds with cooperation to the opponent's cooperation, but in one third of the cases the opponent's defection is answered with cooperation.

In the new paper⁵ Nowak and Sigmund allowed for any possible mutant strategy which, like TFT, remembers only the last move (this includes Einstein's potential proposal if it had a memory of only one move). A strategy was then defined by a set of four probabilities to cooperate, given the outcome of the previous round was R (both cooperated), S (self cooperated, opponent defected), T (self defected, opponent cooperated) or P (both defected), respectively. So, a player's decision depended not only on his opponent's but also on his own last move. TFT would be (1, 0, 1, 0). Each simulation started with the random strategy (0.5, 0.5, 0.5, 0.5). All that happened thereafter was therefore due to mutations. A total of 10^5 different mutant strategies occurred in each of 40 simulation runs — an almost complete imitation of evolution (almost, because of the restricted memory).

After a period of rapid change, long periods of stability alternated with short periods of evolution. Again, defective strategies spread first, followed by a short phase of almost strict TFT. Then either Generous TFT took over or, in more than 80 per cent of the simulations, a new strategy appeared (close to 1, 0, 0, 1) and mostly became stable. This strategy is to cooperate after R and P , and to defect after S and T — in other words, to stay with the previous option after winning the higher payoffs R and T , and shift after S and P . Nowak and Sigmund call it 'win-stay, lose-shift', or 'Pavlov' because of its reflexive nature. Pavlov cooperates when it plays against another Pavlov or any TFT, and it corrects errors quickly. In contrast to TFT it exploits strict cooperators after a while but loses against Always Defect because it

then alternates cooperation and defection. However, a Pavlov population cannot be invaded by Always Defect, if Pavlov is slightly less cooperative after P than after R — for example (0.999, 0.001, 0.001, 0.995) would do.

Which behaviour should empiricists expect to find in real animals? There should be clear differences between populations, but almost no mixtures of strategies within populations. We should see Always Defect or Generous TFT or, most often, Pavlov. Only young populations, with respect to the time they have been exposed to Prisoner's Dilemma selection, are likely to play Always Defect.

These predictions can be tested. Cooperative inspection of predators by small fish⁶ has been investigated both in populations with a long history of predation and in ones where predators have been absent

since records were taken^{7,8}. The fish were indeed much more cooperative at sites with high predation risk. Did they behave more in the manner of Generous TFT or Pavlov? Nowak and Sigmund find cases that have been interpreted as evidence for TFT to be consistent with a Pavlov-like strategy as well, because TFT and Pavlov generate similar predictions in these situations. But conditions can be staged under which TFT and Pavlov diverge. The simplest case would be when the opponent switches from Always Cooperate to Always Defect. This is an animal that defects after it has shown that it can cooperate. Pavlov would switch from cooperation to strictly alternating defection and cooperation, whereas any TFT would not alternate after having switched. In a test of this, my colleagues and I presented stickleback fish with a companion that cooperated during four consecutive runs and thereafter always defected⁹. In looking at those fish which cooperated in the second run after the opponent's first defection, we found that thereafter they indeed alternated defection and cooperation (unpublished data) — here is a glimpse of Pavlov's identity.

'Win-stay, lose-shift' is a rule that is commonly used in situations other than the Prisoner's Dilemma. Experimental psychologists usually find that animals in Skinner boxes repeat rewarded behaviour and drop unrewarded actions. When foraging, animals stay longer in patches that offer a good meal than in ones where food is scarce. After fights over a territory or a harem the winner normally stays and the loser leaves. It is therefore possible that Pavlov has already been evolved when a population is suddenly exposed to Prisoner's Dilemma selection. Only the step of using Pavlov in the new context has to be made. Empiricists should start by testing specifically how animals use their memory when they are in an iterated Prisoner's Dilemma.

The theorists have done a major part of their job in exploring the mysteries of the Prisoner's Dilemma. Now it is the empiricists' turn to find out whether natural evolution has come up with TFT, Pavlov — or anything else. □

Manfred Milinski is in the Abteilung Verhaltensökologie, Department of Zoology, University of Bern, CH-3032 Hinterkappelen, Switzerland.

1. Axelrod, R. *The Evolution of Cooperation* (Basic Books, New York, 1984).
2. Axelrod, R. & Hamilton, W. D. *Science* **211**, 1390–1396 (1981).
3. Nowak, M. A. & Sigmund, K. *Nature* **355**, 250–253 (1992).
4. Nowak, M. A. & May, R. M. *Nature* **359**, 826–829 (1992).
5. Nowak, M. & Sigmund, K. *Nature* **364**, 56–58 (1993).
6. Milinski, M. *Nature* **325**, 433–435 (1987).
7. Magurran, A. E. *Behav. Ecol. Sociobiol.* **19**, 267–273 (1986).
8. Dugatkin, L. A. & Alfieri, M. *Evol. Ecol.* **6**, 519 (1992).
9. Milinski, M., Külling, D. & Kettler, R. *Behav. Ecol.* **1**, 7 (1990).

The end of the last theorem?

Peter Swinnerton-Dyer

ALMOST every day someone somewhere in the world claims to have proved Fermat's last theorem. Invariably those claims have been groundless. As a consequence, the French Academy of Sciences decided some years ago that it would no longer consider purported proofs for publication. The German number theorist E. M. Landau, a natural target for pur-



Pierre de Fermat (1601–1665). Fermat's formal training and career were in law; he essentially studied mathematics only in his spare time, making his achievements all the more remarkable. Among them was his work on the theory of numbers, picking up where the Greek algebraist Diophantus had left off some 1,400 years earlier. It was in the course of that endeavour that Fermat formulated his last theorem — that, where n is an integer greater than 2, the equation $x^n + y^n = z^n$ has no solution in positive integers.

ported proofs, used a printed card thanking his correspondents for consulting him and stating that the "first error" appeared on "line . . . , page . . .". Research students were given as an exercise the task of filling in the blanks.

Even so, the proof of Fermat's last theorem is the most glamorous problem in mathematics, and the game of solving it has been played since the mid-seventeenth century when the theorem was first formulated.

Now it seems that the game is up. In three lectures given at the Isaac Newton Institute, Cambridge, United Kingdom, last week, Andrew Wiles of Princeton University revealed a proof that has all the hallmarks of authenticity. Wiles (who is

British) is a world-class number theorist and a cautious man, and the experts in the audience were confident that his proof is correct.

Fermat's last theorem is the assertion that for any $n > 2$ the only integer solutions of

$$x^n + y^n = z^n$$

are the trivial ones for which at least one of x , y and z is zero. Around 1637, Fermat noted this statement in the margin of his copy of a book on Diophantus's arithmetic and went on to say that he had found a marvellous proof of the result, which unfortunately the margin was too narrow to contain. He had already published proof of the special cases $n = 3$ and $n = 4$, but no one now supposes that he had a valid proof of the general result. If one assumes that there is unique factorization into primes in the field of n th roots of unity (which is unfortunately false) one can derive Fermat's last theorem by arguments very much in Fermat's style. It seems likely that this is what he did.

Until recently, a major obstacle to proving the theorem was that it was an isolated result and no one knew how to attack it. Indeed when, a century or so after Fermat's death, Gauss was asked why he had never attempted to prove the theorem, he replied that he could propose a hundred such propositions which could neither be proved nor disproved and which served only to impede the progress of mathematics. As it turns out, that was an unfair judgement; for the origin of algebraic number theory was Kummer's attempt to provide a satisfactory substitute for unique factorization in the field of n th roots of unity, in the hope that this would lead to a proof of the theorem. Those who built on Kummer's ideas did prove the theorem for many values of n , but it no longer seems likely that those methods will lead to a complete proof.

A completely new line of attack was opened in the 1980s by Ken Ribet, building on a remark of Serre. By considering curves of the form

$$Y^2 = X(X - x^n)(X + y^n)$$

with conductor xyz , where x , y , z is a solution of the Fermat equation, he was able to show that Fermat's last theorem would follow from the Weil–Taniyama conjecture — that any elliptic curve defined over the rationals can be parametrized by elliptic modular functions. That may sound a curiously *ad hoc* statement; but suitably restated it is a key problem in the 'Langlands programme', which when completed will provide a unifying framework for algebraic geometry, algebraic number theory, representation theory, the theory of automorphic func-

tions and maybe more besides.

Thus Fermat's challenge to succeeding centuries will prove to have been much more than an elaborate brain-teaser. In itself, it has been a stimulant of efforts such as those of Kummer, which in turn have led to new conceptual structures. It will be a fitting outcome of the search for a proof if it now emerges that the lasting outcome is a conceptual unification of such disparate fields.

Andrew Wiles has not yet proved the full Weil–Taniyama conjecture, though

that proof cannot be more than a few years off. But he has proved enough of it to be able to deduce Fermat's last theorem. His argument draws on results across almost the whole spectrum of modern pure mathematics. What was announced last week is far from being the end of the story. □

Peter Swinnerton-Dyer is at the Isaac Newton Institute for Mathematical Sciences, 20 Clarkson Road, Cambridge CB3 0EH, UK.

CONSERVATION BIOLOGY

A helping hand in succession

Peter D. Moore

CONSERVATION is in part becoming an exercise in global gardening, and the rehabilitation of ecosystems when the natural vegetation has become highly fragmented is being investigated from many angles. One imaginative strategy is currently being tried in two separate experiments on Staten Island, New York, and in central Florida, the results of both of which are reported in the latest issue of *Conservation Biology*. Robinson and Handel describe their study of the rehabilitation of landfill sites on Staten Island¹ and McClanahan and Wolfe discuss their approach to the reclamation of an abandoned phosphate mining site in central Florida². Both teams have based their work on the idea that the process of succession, particularly the invasion of new species, can be accelerated if the appropriate treatment is applied.

Human land use, particularly urbanization, often has the effect of fragmenting natural habitats, and if any of those fragments are damaged their recovery may be delayed because reinvasion of any locally lost species is often slow as a result of isolation. Similarly, the development of vegetation on isolated reclaimed sites may be limited by the slow arrival rate of colonists. The establishment of new species is related inversely to the distance from any source of new organisms, directly to the area of the site and hence to its potential as a target for invaders, and also directly to the range of microhabitats and microclimates available at the site which may contribute to the successful establishment of new arrivals.

These criteria are essentially those developed in the study of island biogeographic theory by MacArthur and Wilson³, and research such as that of Crowe⁴ on the invasion of weeds into parking lots in Chicago shows that the general principles of the theory do indeed apply to urban 'islands' and reclamation sites. The speed of equilibration of such ecosystem fragments could theoretically

be increased if the immigration rate of potential settlers were increased. This could be achieved entirely artificially by introducing organisms, or by the somewhat more natural technique of encouraging visits by those birds that operate as vectors of the plant species involved in natural succession.

A kind of hybrid approach, adopted in the landfill sites examined by Robinson and Handel, is the planting of trees and shrubs that may subsequently attract birds from other sites that carry the seeds of immigrant plants in their guts. A 1.5-hectare landfill site on Staten Island was covered with a 40-cm cap of compacted subsoil followed by sandy material and organic mulch. Three different vegetation types were then established by planting — an oak–shrub mix, a pine–shrub mix and an ericaceous cover. Further invasion of plant species was recorded within each of the three plantation types, particular attention being paid to species invading from other natural vegetation fragments in the area. In total 18 species had been deliberately introduced, but within a year this number had increased to 50. Nine of the 32 new recruits were wind-dispersed species and 20 were mammal or bird dispersed. The remaining three could have been brought to the site accidentally in the root balls of the planted species.

The arrival of bird-dispersed species depends upon the avian vector finding an appropriate place to perch and to shelter, which suggests that plantations with a higher structural complexity (that is, denser and taller trees and shrubs) should prove more attractive to them⁵. Robinson and Handel tested the density of avian-dispersed species against both the density and the height of the planted species and found a significant positive relationship in both cases. Dense planting of trees and shrubs that rapidly grow tall is evidently an effective means of increasing immigration rates of other plants, and in this way stimulates succession and thus general

vegetation development and recovery.

If complexity of structure is one clue to encouraging plant immigration by the opportunities it offers to visiting birds, then it may be possible to attract the birds without having to wait for the development of vegetation architecture. In other words, if we simply provide perches then birds may well feel inclined to visit a site and defaecate in comfort. This is precisely the approach adopted in the Florida mine reclamation site where McClanahan and Wolfe have conducted their research.

A compacted, sandy soil covered the mine waste after reclamation in 1982. The area was devoid of vegetation but, instead of planting living trees, seven dead trees were erected (mean height about 11 m) and seed traps were positioned both close to them (1–3 m) and at some distance (10 m) to act as a control. A range of frugivorous birds was observed using the perches provided by the dead trees over the following seven years. In a study of seed fallout over a 20-month period, the density of seeds arriving in traps beneath the dead trees averaged 340 m⁻², while the control fallout was only 2 m⁻². The structural complexity of a dead tree is evidently acceptable to visiting birds and provides the necessary enhancement of plant immigration rate.

Actual establishment of species was considerably lower, however. By 1990 only two plants per square metre had become established beneath the dead trees; but this still far exceeded establishment away from the trees, where only 0.07 m⁻² had gained a foothold. A bank of viable seeds (17 m⁻²) was also present beneath the perches, whereas no seed-bank could be detected away from them.

This simple experiment demonstrates very clearly that it is possible for human management to facilitate a succession and short-cut the lengthy process of the development of vegetation canopy. The method is limited in the plant species it can encourage, however, because the changes in soil organic matter, microbiology and hydrology that accompany full vegetation development have not taken place and the residual edaphic rigours may preclude the establishment of all but the most tolerant plant species. But the simple provision of perches places a cheap and useful tool in the hands of the reclamation ecologist — more strength to the arm of the global gardener. □

Peter D. Moore is in the Division of Life Sciences, King's College, Campden Hill Road, London W8 7AH, UK.

1. Robinson, G. R. & Handel, S. N. *Conserv. Biol.* **7**, 271–278 (1993).
2. McClanahan, T. R. & Wolfe, R. W. *Conserv. Biol.* **7**, 279–288 (1993).
3. MacArthur, R. H. & Wilson, E. O. *The Theory of Island Biogeography* (Princeton University Press, 1967).
4. Crowe, T. M. J. *Biogeogr.* **6**, 169–181 (1979).
5. McDonnell, M. J. & Stiles, E. W. *Oecologia* **56**, 109–116 (1983).

Mantle chemistry goes back into the melting pot

Stuart Ross Taylor

THE seemingly innocuous title of the report on page 54 of this issue gives little hint of the shake-up it could mean for earth and planetary sciences. Keppler and Rubie¹ produce evidence that, at high pressure, ions of the transition metals cobalt and nickel tend to change their coordination in silicate melts, becoming octahedrally instead of tetrahedrally coordinated. The upshot is that cobalt and nickel will tend to remain in the melts, rather than enter crystallizing silicates or metal, with profound consequences for the geochemistry of evolving planets.

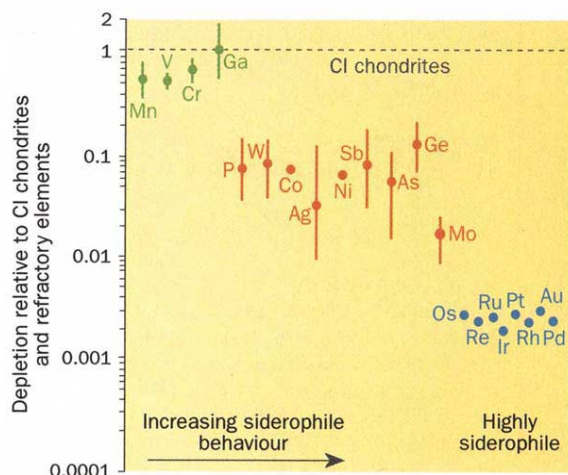
In low-pressure silicate melts, cobalt and nickel both enter early-crystallizing phases such as olivine, one of the main constituents of the Earth's upper mantle. If an iron metal phase is present, as in many meteorites, the cobalt and nickel will prefer that; they are typically present at levels of a few per cent. By analogy with meteorites, nickel and cobalt are widely held to be significant minor components of Earth's core. But meteorites formed at pressures much lower than the megabar pressures in the centre of the Earth.

The distribution of the siderophile elements (those soluble in iron, of which nickel and cobalt are prime examples) in the Earth's silicate mantle has long been a puzzle. As the figure shows, the elements fall into three groups, with decreasing abundance in the silicate mantle correlated with increasing tendency to enter the metallic phase. The step-like pattern formed by these groups, with element abundances equal to roughly 100, 10 and 0.1 per cent of those in chondritic meteorites^{2,3} (the most primitive type), is most striking. The Ni/Co ratio in nodules is often described as chondritic, but it is, in fact, 10 per cent lower⁴. Thus there are minor but perhaps significant deviations from the chondritic abundances in the mantle's siderophile element abundance patterns.

The stepwise abundance pattern comes from analyses of mantle nodules⁵ and of basalts derived from just the upper 200 km of the mantle, a tiny fraction of the 2,900-km thickness of this silicate shell. Much effort has been expended in trying to understand how these abundances relate to the rest of the mantle, and how they are affected by, for example, the recycling of oceanic crust into the mantle and the hot plumes that apparently rise

from the bottom of the mantle.

Taking the upper-mantle abundances at face value, the mantle has too much nickel and cobalt to have been in equilibrium with the metallic core, where, on the basis of the low-pressure elemental distribution coefficients, geochemists would have expected the nickel and cobalt to be concentrated. There is increasing evidence that the Earth was initially molten, either as a



The abundances of siderophile elements in the Earth's upper mantle, normalized to the refractory elements, are generally much lower than in chondrites but are apparently too high to have been in equilibrium with a molten core. The depletion is more extreme for elements of more siderophile character, as this plot shows. The abundances of the volatile siderophile elements have been corrected to allow for the general depletion of the volatile elements in the Earth. (Courtesy of H. E. Newsom, University of New Mexico.)

consequence of the way it grew out of a population of colliding planetesimals, or because of heat generated in a subsequent giant impact widely thought to have created the Moon⁶. If so, we would have expected to find the nickel and cobalt not in the upper mantle, but in the core.

Murthy recently raised the possibility⁷ that the distribution coefficients were affected by the very high temperatures (4,000 K) due to planetesimal accretion and the Moon-forming event, giving calculated values that match the observed distribution. But this work remains controversial⁸; perhaps the real explanation lies in pressure-induced changes in the distribution coefficients. It is this aspect of Keppler and Rubie's work which may force a re-evaluation of the whole question of the distribution of elements with depth in the terrestrial mantle.

A further remarkable geochemical feature of the upper mantle is that its Mg/Si ratio is much higher than that of the solar

nebula. This is sometimes ascribed to flotation of magnesium-rich olivine to upper levels during crystallization of a terrestrial magma ocean⁹. As nickel and cobalt readily enter olivine under low pressure, the upper mantle could be enriched with these elements. But what if olivine is not so favoured? If the whole mantle has this signature and the terrestrial Mg/Si ratio is indeed 'nonchondritic', then the Earth must have been built from a population of planetesimals that also had nonchondritic Mg/Si. This has profound implications for the primordial evolution of the solar nebula — at the very least it means that the Earth's precursors were a very localized suite of planetesimals.

Some caveats remain. In their experiments, Keppler and Rubie used nickel and cobalt-doped albite (NaAlSi₃O₈), as this is readily quenchable to a glass. Would the behaviour of the elements be different in a melt rich in magnesium and iron? This would correspond more closely to the composition of the mantle, where sodium and aluminium are minor components, and where silicon and oxygen tend to form isolated tetrahedra rather than the linked network structure of the feldspars. Perhaps warning should be taken from the dependence of crystal/melt distribution coefficients on the composition of the melt: excursions by a factor of ten are not uncommon. However, taken at face value, Keppler and Rubie's experiments indicate that at high pressure, nickel and cobalt will tend to be stabilized in the melt, rather than entering the silicate or metallic phases.

But what if the Earth accreted from planetesimals already differentiated into silicate mantles and metallic cores? This notion is supported by studies of the asteroid belt¹⁰, where the proportion of differentiated asteroids increases sunwards, reaching 100 per cent at about twice the radius of the Earth's orbit. It

1. Keppler, H. & Rubie, D. C. *Nature* **364**, 54–56 (1993).
2. Newsom, H. E. in *Origin of the Earth* (eds Newsom, H. E. & Jones, J. H.) 273–288 (Oxford Univ. Press, New York, 1990).
3. Arculus, R. J., Holmes, R. D., Powell, R. & Righter, K. in *Origin of the Earth* (eds Newsom, H. E. & Jones, J. H.) 251–271 (Oxford Univ. Press, New York, 1990).
4. McFarlane, E. A. & Drake, M. J. in *Origin of the Earth* (eds Newsom, H. E. & Jones, J. H.) 135–150 (Oxford Univ. Press, New York, 1990).
5. Jagoutz, E. et al. *Proc. Lunar Planet. Sci. Conf.* **10**, 2031–2050 (1979).
6. Benz, W. & Cameron, A. G. W. in *Origin of the Earth* (eds Newsom, H. E. & Jones, J. H.) 61–67 (Oxford Univ. Press, New York, 1990).
7. Murthy, V. R. *Science* **253**, 303 (1991).
8. Jones, J. H., Capobianco, C. J., Drake, M. J. & O'Neill, H. St. C. *Science* **257**, 1281–1285 (1992).
9. Agee, C. & Walker, D. *Earth Planet. Sci. Lett.* **90**, 144–156 (1989).
10. Gaffey, M. J. in *Origin of the Earth* (eds Newsom, H. E. & Jones, J. H.) 17–28 (Oxford Univ. Press, New York, 1990).

RÉSUMÉ

Chicken feed

WHAT, thought J. Pen *et al.*, can be done to avoid the need to supplement the diets of some domestic animals, such as pigs and chickens, with inorganic phosphorus? Feed supplementation with the enzyme phytase, which releases the nutrient from phytate (the main form of phosphorus in many plant seeds), is one way. But Pen and colleagues have tested another (*Bio/Technology* **11**, 811–814; 1993), by engineering the DNA fragment encoding phytase in the fungus *Aspergillus* into tobacco. Transgenic seeds stored the phytase stably, and were effective in releasing inorganic phosphorus in experiments *in vivo*; moreover, poultry whose diet was supplemented with the modified seeds showed significantly higher growth rates than controls. The authors now plan to try the approach with plants whose seeds are more commonly used as feedstuff.

Watch this space

TINY though they are, meteoroids in the Perseid shower could cause considerable damage if they were to strike one of the Earth's many artificial satellites. This year, with the Earth passing close behind the tail of the meteoroid parent body, the infamous comet Swift-Tuttle, the shower may rise to a storm lasting anything from 10 minutes to an hour. Martin Beech and Peter Brown (*Mon. Not. R. astr. Soc.* **262**, L35–L36; 1993) estimate from past records of exceptional Perseid and Leonid showers that an object the size of the Hubble Space Telescope has about a chance in a thousand of being hit by a hefty meteoroid — say, two or three milligrams in mass — this August. Surely no one satellite could be so unlucky?

Molecular poultice

PERHAPS the most disagreeable trick that evolution has taught the malaria parasite is how to induce the infected red blood cell to stick to endothelial surfaces and cause such complications as cerebral malaria, which tends to be fatal. The mechanism of adhesion is therefore a busy area of malaria research. A suggestion now emanating from I. W. Sherman's laboratory is that adhesion is promoted by a modified form of the abundant host cell membrane protein, band 3. From fragmentary evidence that histidine, tyrosine and lysine side chains contribute to adhesion, I. Crandall *et al.* (*Proc. natn. Acad. Sci. U.S.A.* **90**, 4703–4707; 1993) have deduced that two of the small external loops of the band 3 chain (which crosses the membrane 14 times) may be the sites in question. They have made corresponding synthetic peptides and find that these indeed inhibit adhesion *in vitro*, and when injected into infected monkeys cause release of the parasitized cells into the circulation. Therapeutic possibilities are to be explored.

is reasonable to suppose that closer to the Sun, where the terrestrial planets accreted, the planetesimals were similarly differentiated. In these small bodies, the pressures during separation of the elements would be fairly low. The Earth's metallic core would segregate at the time of accretion, with perhaps only limited opportunity for re-equilibration of metal and silicate under high pressures.

Nonetheless, if the Earth was subsequently remelted by a giant Moon-forming impact, or if the silicate mantle remained molten as a consequence of accretion, then the crystallization of the deep silicate mantle would have occurred under high pressure and the results of Keppler and Rubie will apply. With so dramatic a change to our views on the distribution of the siderophile elements in

the mantle, it will be back to the drawing board for the high-pressure geochemists, as 30 years of study have not cast much light on the paradox of elemental abundances in the upper mantle.

The most popular and perhaps the most reasonable notion, that the signature represents the addition of a 'late veneer' of chondritic material is not without its own difficulties, of which the most telling is the absence of a similar veneer on the Moon. Given the present complexity of the mantle, our inadequate sampling and the possible effects of a giant impact, others must follow the lead given by Keppler and Rubie if we are to solve the paradox. □

Stuart Ross Taylor is in the Department of Nuclear Physics, Australian National University, Canberra, ACT 2601, Australia.

IMMUNOLOGY

MHC class II dimer of dimers

Hidde Ploegh and Philippe Benaroch

SIX years have passed since publication¹ of the three-dimensional structure of a class I molecule of the major histocompatibility complex (MHC). On page 33 of this issue² we are now treated to a similar feat, accomplished for a human class II MHC molecule, HLA-DR1 — and by the same laboratories no less.

Unlike immunoglobulins, antigen specific receptors on T lymphocytes (TCR) do not interact with intact, foreign proteins, but recognize short (8–25-residue) peptides derived from them, held tightly in the jaws of MHC molecules. MHC proteins come in two basic types, class I and class II, each of which subserves a special type of T cell. Class I products interact with CD8 (largely cytotoxic) T cells, whereas class II molecules interact with CD4 (predominantly helper) T cells. What is good for the goose is good for the gander, for it turns out that the structures of both types of MHC product are remarkably similar (a strong case for which, incidentally, was made some years ago³). From similarities in primary structure of class I and class II MHC products, Brown *et al.*⁴ had already proposed a model for the class II MHC binding site that has stood workers in the field in good stead: experiments based on the model produced results entirely consistent with it. Its vindication now comes from the structure before us.

Buried in the fine print are references to a struggle lasting almost ten years to obtain enough material in pure enough form⁵ to proceed with attempts at crystallization⁶. But that persistence has paid off: no fewer than three different crystal forms — with the added bonus of a non-crystallographic dimer — could be

used to solve the structure of HLA-DR1. At first glance, a class II molecule is easily mistaken for its class I counterpart. But just as the class I structure drove a stake through the heart of any remaining dual-recognitionist demon by unveiling its peptide cargo, so the class II molecule reveals new secrets.

First is the fundamental difference between class I and class II molecules in peptide-MHC interactions. The class I peptide-binding pocket is blocked at either end and appears to impose severe restrictions on the sizes of peptides it can accommodate (8–10 residues), with longer peptides bulging out in the middle⁷. Not so for class II. The class II binding groove allows peptide to protrude from it and, consequently, longer (average 15–18-residue) peptides⁸ can bind: there is no need for bulges here.

Studies of self-peptides eluted from purified HLA-DR alleles have directly established this size heterogeneity, as well as the occurrence of nested sets of peptides derived from a single protein^{8,9}. Such peptides are thought to bind through a central conserved core with 'ragged' ends. Many of the contacts of DR1 with its bound peptide — in the same amino- to carboxy-terminal orientation as in class I — involve DR1 side-chains and main-chain atoms of the peptide. Only a single 'pocket', capable of accommodating a rather large peptide side chain, is observed. Other pockets are indicated by patches of polymorphic residues in the binding site. In addition, promiscuous peptides, capable of binding to many different human class II alleles, have been identified^{8,10}, with no obvious parallel in class I products.

The mode of binding of peptide to class II molecules may thus be governed by different rules (including those that follow from trafficking, generation of peptides and the like) from those that dictate class I-peptide interactions. This is an important issue, not only for understanding the biology of antigen presentation but also for the purpose of manipulating the immune response. Further refinement of the HLA-DR1 structure, elucidation of other class II structures and the ever-expanding database of naturally processed peptides will tell us more about it.

But the major surprise is that the DR1 heterodimer, in all three crystal forms, occurs as a dimer of dimers — here called $(\text{dimer})_2$ — with an orientation that would allow simultaneous interaction with two TCR complexes. If this finding is generalized to other class II MHC products, it will be difficult to resist the attractive idea that $(\text{dimer})_2$ might exist at the cell surface, but only under certain conditions².

How could a T-cell response be set in motion? Two TCR complexes would act as molecular tweezers that sort out and bring together, on the surface of the antigen-presenting cell, class II heterodimers occupied by the same — or similar — peptides. Two copies of the CD4 molecule can then join the party. The resulting aggregate would have a greater stability than single TCR-class II complexes (which are of low affinity¹¹). Also, each CD4 coreceptor could now interact with both of the two class II heterodimers in $(\text{dimer})_2$. Apart from any conformational change, the physical proximity of all participating subunits¹² (including associated CD3 and attendant kinases) imposed by this aggregate might set off TCR triggering as well as signal transduction through class II molecules. The antigen-presenting cell must be told that it has been contacted by a specific T cell, and signalling by way of $(\text{dimer})_2$ would ideally suit that purpose. Upregulation on the antigen-presenting cell of the costimulatory molecule B7/BB1 would ensue, and its binding to the CD28 and CTLA-4 receptors on the T cell could now deliver the second (costimulatory) signal necessary for T-cell activation¹³.

It is tempting to consider other TCR-mediated phenomena in light of the $(\text{dimer})_2$ model. Antagonist peptides¹⁴, which are analogues of antigenic peptides that can bind class II MHC products and specifically inhibit, rather than stimulate, T-cell activation, could poison half of $(\text{dimer})_2$. In doing so, they might impede proper signal transduction or could result in the delivery of a qualitatively different signal. Alternatively, the lower affinity of specific TCRs for antagonist-loaded molecules could preclude formation of a stable $(\text{dimer})_2$. Furthermore, during thymic development, T cells acquire the ability to interact with the MHC products

of the host (positive selection), while potentially deleterious T cells are removed (negative selection). Positive selection of CD4⁺ T cells, which imposes self-restriction regardless of antigenic specificity and requires recognition of self class II MHC products, might occur without formation of $(\text{dimer})_2$, whereas negative selection, which involves specific recognition of MHC-peptide complexes, might depend on it.

Given the structural similarity between class I and class II molecules, we should entertain the notion that class I molecules display similar dimerization behaviour. The T-cell CD8 co-receptor, being a homodimer that interacts with class I molecules, might indeed bind two of them at once. However, dimers of class I molecules have not been observed in the numerous crystals examined to date, and the necessary contact residues that stabilize the dimer may simply not be there.

As for class II molecules, although biochemical supporting evidence is still

lacking, might it be that $(\text{dimer})_2$ actually triggers not only our imagination but also T cells? Cell lines and mice carrying mutants of class II molecules genetically altered to prevent $(\text{dimer})_2$ formation will probably point the way to an answer. □

Hidde Ploegh and Philippe Benaroch are at the Center for Cancer Research, Massachusetts Institute of Technology, Building E17-322, Cambridge, Massachusetts 02139, USA.

1. Bjorkman, P. J. *et al. Nature* **329**, 506–512 (1987).
2. Brown, J. H. *et al. Nature* **364**, 33–39 (1993).
3. Kaufman, J. F. thesis, Harvard Univ. (1983).
4. Brown, J. H. *et al. Nature* **332**, 845–850 (1988).
5. Gorga, J. C., Horejsi, V., Johnson, D. R., Raghupathy, R. & Strominger, J. L. *J. biol. Chem.* **262**, 16087–16094 (1987).
6. Gorga, J. C., Brown, J. H., Jardetzky, T., Wiley, D. C. & Strominger, J. L. *Res. Immun.* **142**, 401–407 (1991).
7. Guo, H.-C. *et al. Nature* **360**, 364–366 (1992).
8. Chicz, R. M. *et al. J. exp. Med.* (in the press).
9. Chicz, R. M. *et al. Nature* **358**, 764–768 (1992).
10. Sinigaglia, F. *et al. Nature* **336**, 778–780 (1988).
11. Matsui, K. *et al. Science* **254**, 1788–1791 (1991).
12. Weiss, A. *Cell* **73**, 209–212 (1993).
13. Schwartz, R. H. *Cell* **71**, 1065–1068 (1992).
14. De Magistris, M. T. *et al. Cell* **68**, 625–634 (1992).

OPTOELECTRONICS

Marvels of molecular device

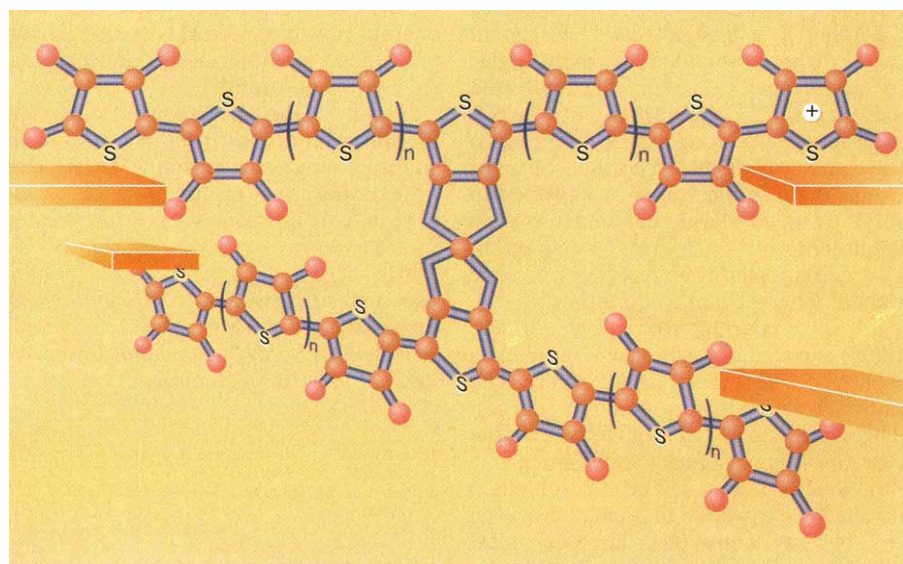
Leonard F. Lindoy

As electronic components shrink, so the interest in molecular-scale devices grows. A step towards producing such devices is reported on page 42 of this issue by de Silva and colleagues¹, who describe how an anthracene derivative can be persuaded to act as a chemical 'AND' logic gate.

What motivates the current interest in molecular-scale devices? With smaller size there is a tendency towards faster

performance, as conducting paths are shorter. Also, a greater concentration of function per unit area in a two-dimensional array should aid the development of efficient parallel-processing devices. These achieve greater speed by allowing signals to travel simultaneously by different pathways through the device, rather than sequentially as occurs in conventional integrated circuits.

Simple, single-input devices could be



A proposed molecular switch, showing electrode connections to the two crossed arms of a spiro molecule. The two switching electrodes for applying the electric field are not shown.

said to have been demonstrated before. A molecule that fluoresces only when well supplied with protons acts, in effect, as a 'YES' gate; conversely, if protons inhibit the fluorescence, we have a 'NOT' gate. The new molecule is more complex. It contains two recognition sites, corresponding to two input channels: the first site binds hydrogen ions, and the second (crown ether) site binds sodium ions. Variation in fluorescence — the output — depends on whether the molecule is provided with hydrogen ions, sodium ions or both. The properties of the dual input/single output arrangement in this system correspond to an AND gate, as strong fluorescence is seen only when both types of ion are present.

Although potentially useful as they stand for ion monitoring, these single-molecule systems provide only rudimentary models for future logic devices, for they operate in solution. Solid-state devices would be far more convenient, although one of the problems then is how to communicate with the device from the outside world; in other words, how to feed in the input and read the output. Some type of circuitry will be necessary, perhaps in the form of conducting molecules to act as molecular 'wires'. It might be possible to communicate with particular devices by applying a magnetic or electric field, or by monitoring the change in absorption or transmission of light. Communication remains one of the chief challenges associated with this field.

One approach is illustrated by the molecular system proposed by Aviram². In developing a blueprint for a molecular switch, Aviram based his designs on the 'spiro' molecule, incorporating two orthogonally oriented thiophene strings (see figure). The thiophene chains are associated with extensive π -electron delocalization and become conductors on oxidation (that is, on electron removal). The chains are linked by a 'bridge' of two pentagonal carbon rings (σ -bonded) at right angles to each other. Overall, the molecule amounts to a π - σ - π arrangement with the saturated spiro bridge corresponding to the σ component. In the normal course of events the σ bridge might be expected to serve as an insulator, but under certain conditions it can permit conduction, allowing transfer of an electron from one polythiophene chain to the other.

In the oxidized form, one of the thiophene chains contains one more electron than the other (in the figure, the latter chain bears the positive charge). This arrangement is equivalent to the electron being in a box with the shape of a double-well potential. The electron might normally be expected to oscillate between the two wells by quantum-mechanical tunnelling. But the incorporation of the spiral moiety makes the two possible states orthogonal and theory indicates

that this will inhibit tunnelling. Molecular orbital calculations predict that applying an electric field (directed along the axis that passes through the central spiro carbon and the two thiophene rings adjacent to the bridge) will alter the relative energies of the two states; if the field is strong enough, electron transfer or 'switching' from one chain to the other may occur³.

Already there is progress towards the practical synthesis of suitable spiro-bridged polythiophene molecules⁴, and modern developments in nanolithography now make it possible to fabricate patterns with line dimensions down to about 15 nanometres, to act as the required connectors and accompanying circuit.

Difficulties still remain. The construction of a successful device of this type requires the thiophene chains to assemble, or be assembled, across the appropriate gaps in the 'wiring' such that their ends make electrical contact with the wiring. It is likely that suitable reactive chemical groups at the terminus of each chain will be needed to make such connections. For gold wiring, thiol groups may prove suitable; their affinity for binding to gold surfaces has been well established⁵.

The above examples are just some of the many approaches being tried to overcome the formidable difficulties in producing practical molecular electronic and optoelectronic components. Uncertainties abound. Two-dimensional arrays are not necessarily the ideal for future devices. Three-dimensional ordered arrays of molecules, with, for example, 'tunable' intermolecular orbital overlaps in different directions (which could be used to control electron movement in each direction), may also prove important, especially if the system is able to self-assemble. It is not even clear that the fast speeds associated with optical and/or electron processes will necessarily dominate, as the use of parallel processing may enable slower processes (such as ion movement or molecular conformational change) to be useful.

Despite these uncertainties, engineering and scientific advances in critical areas continue across a wide front. There seems little doubt that useful molecular-scale devices will be made in the foreseeable future. □

Leonard F. Lindoy is in the Department of Chemistry and Biochemistry, James Cook University of North Queensland, Townsville 4811, Queensland, Australia.

1. de Silva, A. P., Nimal Guaratne, H. Q. & McCoy, C. P. *Nature* **364**, 42–44 (1993).

2. Aviram, A. *Int. J. Quantum Chem.* **42**, 1615–1624 (1992).

3. Hush, N. S., Wong, A. T., Bacskaý, G. B. & Reimers, J. R. *J. Am. chem. Soc.* **112**, 4192–4197 (1990).

4. Tour, J. M., Wu, R. & Schumm, J. S. *J. Am. chem. Soc.* **112**, 5662–5663 (1990).

5. Labinis, P. E., Hickman, J. J., Wrighton, M. S. & Whitesides, G. M. *Science* **245**, 845 (1989).

Bars to entry

DREADCO has a company cat: a friendly tabby, which enters the premises through its own cat-flap. To stop its friends coming in as well, Daedalus has fitted the cat-flap with a supermarket bar-code reader. It reads every cat approaching the flap, and lets it in only if it shows the correct pattern of stripes.

Daedalus now wants to use this scheme on the company's human employees as well. Current card and key-based security systems are awkward, annoying and often go wrong. But a bar-code reader operates automatically, works at a distance, and can read a code presented at any speed, in either direction, and at a wide range of angles. Properly sited, it could scan and identify everyone who enters or leaves the building.

Even Daedalus hesitates to brand or tattoo bar-codes on his work-force. But many traditional garments, from prisoners' uniforms to pin-striped trousers, have a strong striped motif already. So DREADCO tailors are devising a machine-readable company dress code. If it works, they plan to licence it to other organizations as well. They are developing machine-readable striped trousers for the British civil service, digitally striped shirts for the City, and coded variants of striped and barred rank-insignia for the military.

Scottish organizations could exploit the bar-code compatibility of the tartan, which after all was invented to distinguish one clan from another. Its vertical and horizontal stripes have a further advantage: an inebriated wearer could crawl horizontally into the premises and still be recognized.

The final system will classify the employees as well as identifying them. Among civil servants, for example, permanent secretaries would need a different pattern of pin-stripe from administrative trainees. Differentially coded patterned tights could be designed for female bureaucrats. A security tailoring company might have to change the patterns regularly, to frustrate intruders who try to sneak into the hallowed halls of authority in forged tights or trousers.

As a simplistic first step, however, Daedalus is scanning his motley work-force with a bar-code reader anyway. Anybody walking past it will generate some code number, simply by his pattern of light and shade. With luck this code will be quite characteristic. Once all the permitted staff have been coded in this way, the system may spot nearly all intruders. They couldn't even disguise themselves as legitimate employees: they wouldn't know what features to imitate.

David Jones

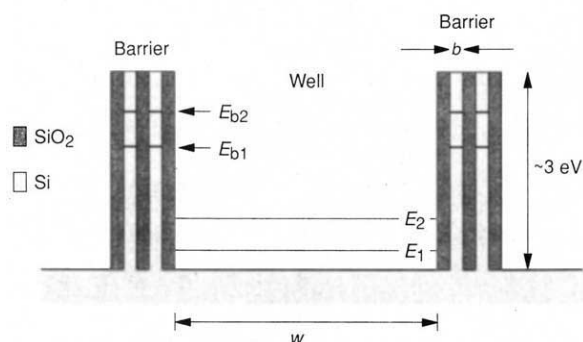
Silicon-based quantum wells

SIR — Since our proposal¹ and subsequent observation² of resonant tunnelling in quantum-well structures, the field has expanded to include all kinds of quantum³ and functional devices⁴. As long as the electron mean-free-path is greater than the dimensions of the quantum well, pseudo-eigenstates can form in the well. To date, the wells and barriers are formed with lattice-matched systems such as

based quantum well.

It is often assumed in strain-layer superlattices that a large lattice mismatch results in high defect density which destroys quantum confinement. But in fact, not only is scattering less effective in lower-dimensional systems, but it fails to destroy phase coherence in zero-dimensional systems such as quantum dots if the scattering process possesses time-reversal symmetry⁹. Consequently, in an isolated system such as a quantum well, defects may not be detrimental.

A strain-layer epitaxial Si/SiO₂ barrier system is shown in the figure. Because the width $b \ll w$, E_{b1} and E_{b2} are much higher than E_1 and E_2 , the quantum well states. This is the basis of a strain-layer barrier. The effective barrier height may be as high as 3 eV, so it is possible to design the separation $E_2 - E_1$,



A Si/SiO₂ quantum well using a Si/SiO₂ superlattice as the barrier.

GaAs/AlGaAs, GaInAs/AlInAs and so on. Because the electronics industry is overwhelmingly dominated by silicon integrated-circuit processing technology, these new devices cannot play an economic technological role unless quantum confinement in silicon can be achieved with barriers much higher than those possible with Si/Si_{1-x}Ge_x systems⁵. Without a higher barrier, it is impossible to create a device operating at a voltage corresponding to an energy much above $k_B T$ at room temperatures. Because of the lack of a suitable material for barriers, I propose that strain-layer superlattices⁶⁻⁸ be used as barriers.

Basically, the concept of a strain-layer superlattice is that, with a sufficiently thin epitaxial layer, the strain energy in each layer is below the energy needed for the growth of point defects or dislocations. It is important to note that dislocations have an activation energy for nucleation and a lower activation energy for growth. Therefore, in principle it is possible greatly to exceed the energy requirement without actually generating defects.

The MOS device owes its success to the low defect density at the Si/amorphous-SiO₂ interface. With a barrier height of 3.2 eV, amorphous SiO₂ should be an ideal barrier for quantum confinement in silicon. Unfortunately, it is not possible to grow epitaxial silicon on amorphous SiO₂. But if the SiO₂ layer is very thin (a couple of monolayers) it remains crystalline, so it should then be possible to grow on top of it an epi-layer of silicon. This is the basis for my proposal that a strain-layer superlattice be used as the barrier for a silicon-

depending on the width w , to be much greater than $k_B T$ at room temperature. If, for example, one can generate four states with separations of ~ 0.5 eV, a four-level multiple logic circuit operating at steps of 0.5 V could be designed.

Other materials for our scheme could be the Si/Al₂O₃ system, which has been known for many years, or the more recent Si/CaF₂ system¹⁰. My choice is the Si/SiO₂

system because SiO₂ can be grown by exposing a silicon substrate heated to ~ 550 °C to oxygen. Subsequently, silicon can be evaporated or even deposited epitaxially with a molecular source. Hydrogen implantation followed by moderate annealing could be used to provide hydrogen passivation of possible dangling bonds. Therefore it may be better to deposit the SiO₂ onto a silicon surface oriented several degrees off any high symmetry axis, such as the (100), to avoid nucleation of defects at the steps.

Finally, to take advantage of small devices (low RC discharge time), it is desirable further to confine the electrons into a dot. Ultimately, the full use of such a quantum-well device will rely on nanoscale lithography.

Raphael Tsu

Department of Electrical Engineering,
University of North Carolina,
Charlotte, North Carolina
and
Nanodynamics, Inc.,
510 E 73rd Street, New York,
New York 10021, USA

1. Tsu, R. & Esaki, L. *Appl. phys. Lett.* **22**, 562–564 (1973).
2. Chang, L. L., Esaki, L. & Tsu, R. *Appl. phys. Lett.* **24**, 593–595 (1974).
3. Capasso, F. *Physics of Quantum Devices*, 181–248 (Springer, Berlin, 1990).
4. Capasso, F. & Kiehl, R. H. *J. appl. Phys.* **58**, 1366–1368 (1985).
5. Chen, Y. F. *et al. Appl. phys. Lett.* **62**, 2713–2715 (1993).
6. Tsu, R. US Patent No. 5216262 (1993).
7. Matthews, J. W. & Blakeslee, A. E. *J. cryst. Growth* **32**, 265–273 (1976).
8. Osbourn, G. C. *J. appl. Phys.* **53**, 1586–1589 (1982).
9. Tsu, R. *Physica B* (in the press).
10. Sinharoy, S., Gregg, J. & Schmidt, D. N. *J. appl. Phys.* **64**, 6296–6300 (1988).

Database of ancient sequences

SIR — Green *et al.*¹ have introduced the concept of ancient conserved regions (ACRs), defined as contiguous amino-acid sequence segments predating the coelomate radiation 500–600 million years ago. Statistical arguments strongly suggest that currently known proteins of vertebrates, invertebrates, yeast and bacteria may already include representatives of most ACRs. Thus the numerous new

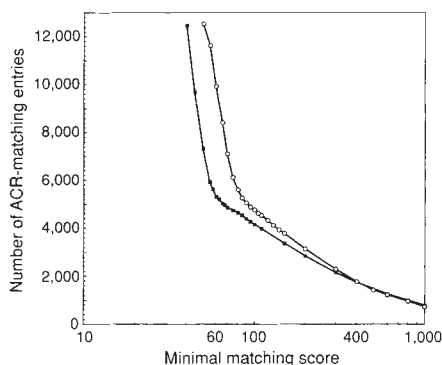
protein domains or motifs that will emerge from sequencing yeast or nematode are likely to be restricted to these phyla.

I have estimated the total number and assembled a repertoire of these ancestral sequences from an analysis of the Swiss-Prot (21.0) protein database². First, I have isolated seven comprehensive subsets of protein sequences from eubacteria, plants, fungi, slime mould, vertebrates,

CROSS-PHYLUM COMPARISONS OF SELECTED SEQUENCE SUBSETS

Subsets	Entries	Target entries	Score > 84 ($P = 5 \times 10^{-5}$)		Score > 69 ($P = 2.5 \times 10^{-2}$)	
			Matching other phyla	ACRs	Matching other phyla	ACRs
Vertebrates	6,578	6,410	2,323 (35%)	390	2,605 (40%)	374
Bacteria	2,543	10,445	521 (20%)	281	602 (24%)	280
Plants	2,224	10,762	860 (39%)	199	975 (44%)	210
Yeasts	984	12,004	561 (57%)	303	606 (62%)	307
<i>Drosophila</i>	440	12,548	317 (72%)	153	334 (76%)	139
Slime mould	124	12,864	68 (54%)	45	73 (59%)	47
Nematode	94	12,894	68 (72%)	45	69 (73%)	40
All	12,988	NA	4718 (36%)	551	5,264 (40%)	520

The final set of ACR representatives is obtained by comparing the ACRs defined from each subset, forming clusters of related sequences and retaining a single sequence by similarity cluster. Although the number of cross-phyllum matches varies with the score threshold (see figure), the final number of independent ACR representatives remains quite stable. Matches were scored according to the optimal pam120 matrix. Change in the scoring matrix did not significantly alter those results.



Number of database sequences matching the 551-ACR subset at various score thresholds. Two different scoring matrices, pam120 (squares) and pam250 (circles), produce similar results. Scores are on a logarithmic scale. The distribution is limited on the right (high scores) by the 551 ACRs matching themselves and on the left (lower score) by the random matching of unrelated sequences. Both curves exhibit a quasi-linear part corresponding to the progressive recruiting of significant matches. An abrupt change in the slope occurs at $S=60$ (pam120, $P \leq 2-5 \times 10^{-2}$) or $S=85$, pam250, $P \leq 2-5 \times 10^{-2}$ where random matches begin to occur. Up to 50% of the entries could be ACR-related.

arthropods and nematodes. For each subset, I compared every sequence (using the local alignment program Blastp³) to the other sequence subsets, recording all matches scoring above a significant threshold (see table).

I then submitted the resulting set of cross-phylum matching sequences (ACR-containing sequences) to self-comparison (again with Blastp and the same score threshold), and partitioned them into clusters of related sequences. For each similarity cluster, I then selected as a representative the member exhibiting the largest conserved region (computed in terms of total length of the cross-phylum matching segment(s)).

This protocol gave an ACR representative set of 551 independent sequences (see table). This result is remarkably insensitive to the details of the procedure (scoring matrices, significance threshold, and so on). Half the 551 ACR representatives correspond to enzymatic activities representing all basic cellular systems: DNA replication, transcription, translation, signal transduction, glycolysis, tricarboxylic acid cycle and oxido-reduction. Structural elements such as ribosomal proteins, histones, cytoskeleton and matrix components are all present. But about 200 ACR representatives correspond to sequences of less obviously central importance, and warrant further analyses.

The 551-representative set constitutes 1,061 distinct contiguous segments with lengths ranging from 18 to 1,497 residues, with an average of 131 residues and a median of 84. Those numbers in part depend on the local nature of the Blastp algorithm, and do not accurately reflect

the domain organization of the ancestral proteins.

It has previously been suggested that proteins are put together from a limited set of 1,000–7,000 exons⁴, or 1,000 distinct folding domains⁵. The coincidence of these numbers is striking. But ACRs are too variable in size simply to be interpreted as structural domains. Furthermore, more than 96% of the database sequences related to the ACR set match a single representative, suggesting that these ancestral segments are rarely independently shuffled in contemporary sequences. Remarkably, 40–60% of the sequences in the database appear related to the 551-ACR representative set (see figure). This small 'ancestral' subset thus constitutes a convenient resource for the fast screening and identification of new sequences (for instance numerous cDNA partial sequences) and the definition of motifs. The 551-representative ACR set is available on e-mail request (jmc@ncbi.nlm.nih.gov).

J.-M. Claverie

National Center for Biotechnology
Information,
National Library of Medicine,
NIH, Bethesda, Maryland 20894, USA

1. Green, P. *et al. Science* **259**, 1711–1716 (1993).
2. Baïroch, A. *Nucleic Acids Res.* **20**, 2019–2022 (1992).
3. Altschul, S. F. *et al. J. molec. Biol.* **215**, 403–410 (1990).
4. Dorit, R. L., Schoenbach, L. & Gilbert, W. *Science* **250**, 1377–1382 (1990).
5. Chothia, C. *Nature* **357**, 543–544 (1992).

Species concepts

SIR — The biological species concept (BSC) is a valuable heuristic that rightly focuses attention on the fundamental importance of reproductive and genetic connections among populations. Nevertheless, we disagree with Mayr's notion¹, discussed by Diamond in News and Views², that the BSC applies easily to plants, and argue that those phenomena that provide the most difficulty for the BSC (for example asexual reproduction, hybridization, polyploidy) represent critical aspects of plant evolution rather than rare aberrations in a flora consisting mostly of a 'good species'. We do not intend to belittle the BSC, nor to suggest an alternative, but argue that the unavoidable limitations of the BSC indicate the need for different species concepts for different taxa, so that the concept does not obscure the variety of biological processes.

For 83 per cent of the species considered by Mayr¹, the typological species concept gave unambiguous and accurate identification. Use of the BSC helps for another 10 per cent, but other species concepts (such as the evolutionary species concept) would probably do just as well for such a broad range of taxa (from

aspen to clubmosses). Indeed, no account is taken of how often ambiguous results might conceivably occur in such a diverse flora (for example by focusing on congeners, for which hybridization is more likely). Comparison of the success of different species concepts would in any case be more valuable than arguing that 93 per cent is an acceptable criterion of success. Unless the usefulness of other species concepts (such as the evolutionary³, phylogenetic⁴ or typological⁵ species concepts) in dealing with these taxa is also assessed, there is no valid comparison on which to base acceptance or rejection of a species concept.

Furthermore, the flora studied by Mayr is a small temperate one that includes few members of the groups that give species concepts the most difficulty, such as orchids. Mayr's survey also does not address the performance of the BSC in describing geographical variation within a plant taxon.

Use of a particular species concept always depends on the questions considered. Those interested in speciation and reproductive isolation might not abandon the BSC even if another species concept helped resolve more difficulties, because the BSC embodies a rich background of evolutionary theory dealing with barriers to hybridization. Similarly, many plant species are asexual or apomictic, and well over a third have hybrid origins⁶, so botanists are understandably reluctant to abandon consideration of these factors when delimiting species.

The issue is not how universally a species concept applies to plants or animals, but the implications of the obvious difficulties in its application for some taxa. Even a single species that clearly does not fit the BSC leads an evolutionarily trained botanist to wonder if modes of speciation other than the traditional route of allopatry are important. We suggest that those instances yield insights into the speciation process for both plants and animals.

**Randall Mitchell, David Bleakly,
Robert Cabin, Raymund Chan,
Brian Enquist, Ann Evans,
Timothy Lowrey, Diane Marshall,
Steven Reed, George Stevens**
Department of Biology,
University of New Mexico,
Albuquerque, New Mexico 87131, USA
Nickolas Waser
Department of Biology,
University of California,
Riverside, California 92521, USA

1. Mayr, E. *Am. J. Bot.* **79**, 222–238 (1992).
2. Diamond, J. *Nature* **360**, 627–628 (1992).
3. Ashlock, P. D. *Syst. Zool.* **28**, 441–450 (1979).
4. de Queiroz, K. & Donoghue, M. *Cladistics* **4**, 317–338 (1988).
5. Mayr, E. & Ashlock, P. D. *Principles of Systematic Zoology* (McGraw Hill, New York, 1991).
6. Riesberg, L. H. & Soltis, D. E. *Evol. Trends Plants* **5**, 65–81 (1991).

Enigmas in the making

Ian Stewart

Collected Works of A. M. Turing. 3 Volumes*. By A. M. Turing. North-Holland: 1992/1993. Distributed by Elsevier.

ALAN Mathison Turing is one of the great original minds of the twentieth century, a founding father of the electronic computer whose name lives on in such concepts as the Turing machine, a simple model for the computational process, and the Turing test, a method he proposed for deciding whether a machine is engaged in genuine thought. This handsome edition of his collected works — of which a fourth volume, *Mathematical Logic*, has yet to appear — includes much unpublished material, and offers new insights into his interests and his mentality. As well as being a great scientist, Turing was also a tragic figure, a quiet homosexual who was pushed to suicide by a discriminatory judicial system. But it is his scientific work that we encounter here.

Turing published very little: in these three volumes we find one published paper on morphogenesis, eight on pure mathematics and seven on machine intelligence (of which two are technical reports and one is a 1947 lecture that has only recently seen print). Perhaps half of the material consists of drafts of papers and articles left in Turing's archives.

The seminal paper *The Chemical Basis of Morphogenesis* dates from 1952. Turing aims to model the developing embryo in terms of the diffusion and reaction of chemical substances called morphogens. Among his key ideas is the spontaneous formation of patterns arising from instabilities of the homogeneous state. He anticipates, in this special case, many modern ideas about symmetry-breaking and bifurcation. The specific mechanism that he proposes did not stand up to detailed experimental scrutiny, and has not found much favour among mainstream biologists, especially in an age that is strongly oriented towards the molecular view of biology and the DNA code. But Turing's central idea has been modified by several groups of mathematical biologists to make it more compatible with known biological facts, and there is a growing body of evidence that the development of biological form must involve dynamic, as well as molecular, processes. In this volume we also find unpublished papers on phyllotaxis — the arrangement of leaves on plants.

Turing's work in pure mathematics is

focused on three main areas: group theory, analytic number theory and the word problem. A paper of 1938 provides a partial answer to a question of Stanislaw Ulam on the approximation of groups by finite groups; the proof is rather interesting, involving an interplay between the representation theories of fi-

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Alan Turing (1912–1954).

nite and compact Lie groups. Turing's number-theoretic research centres on the Riemann zeta function, whose analytic properties are important for the distribution of prime numbers. Perhaps the unsolved problem in mathematics is the Riemann hypothesis that all non-real zeros of the zeta function lie on the line 'real part = $1/2$ ' in the complex plane. Turing put a lot of effort into devising effective numerical methods for locating these zeros accurately: here we see his interest in computers impinging on other, more ethereal areas. A quite remarkable unpublished manuscript, written jointly (or perhaps not) with S. Skewes, deals with the error term in the prime number theorem. This theorem relates the number $\pi(x)$ of primes less than some bound x to the logarithmic integral $\text{li}(x)$. Carl Friedrich Gauss conjectured, and Jacques Hadamard and Charles-Jean de la Vallée Poussin proved, that the ratio $\pi(x)/\text{li}(x)$ tends to 1 as x tends to infinity. Numerical

evidence available in the early 1900s suggested that $\text{li}(x)$ is always greater than $\pi(x)$, but John Edensor Littlewood proved that the difference changes sign arbitrarily infinitely often. Skewes, his student, proved that $\text{li}(x)$ is less than $\pi(x)$ for some x less than the gigantic number $10 \uparrow 10 \uparrow 10 \uparrow 34$ (where $\uparrow$ is exponentiation), on the assumption of the Riemann hypothesis. Skewes and Turing attempt to improve this to $\exp(\exp(661))$, without assuming the Riemann hypothesis. Their method is basically sound but some errors need repairing, and the value of x becomes $\exp(\exp(1236))$, as A. M. Cohen and M. J. E. Mayhew reported in 1965. Their paper is included.

The most important paper here is a proof that the word problem in cancellation semigroups is insoluble. The gist of the problem is that an algebraic structure of a particular kind is defined by sequences of symbols, together with transformation rules; we then ask whether an algorithm exists to decide when two symbol strings are equivalent under such transformations. Turing proved that the answer is 'no', by relating the problem to his work on Turing machines. Subsequently, P. S. Novikov proved the insolubility of the word problem for groups, building on Turing's result.

Finally, we have the work on machine intelligence. In a paper of 1948, Turing predicts three important ideas: sub-routines, automated software verification and machine learning. The key offering is the 1950 paper in *Mind*, which begins: "I propose to consider the question, 'Can machines think?'" Turing's proposed test for machine intelligence is hotly debated even today; the original paper is well worth re-reading, although the central issue is more one of definition — influenced by ideology — than of philosophy.

Turing was a deep thinker whose career was hugely disturbed by the Second World War and who in any case felt little urge to publish. Accordingly, the overall effect of his collected writings is rather bitty — an impression not entirely helped by the format of these volumes, with their many prefatory notes, analyses and interpolated explanations. But the sheer power of the man's mind shines through like a beacon. It is a pity that the hypocrisy of a bygone age put a premature end to such a sparkling scientific career; but we should not be too smug about today's more relaxed attitudes when the careers of many equally promising young scientists are being snuffed out, even before they start, by a political system that knows the price of everything and the value of nothing. □

Ian Stewart is in the Mathematics Institute, University of Warwick, Warwick, Coventry CV4 7AL, UK.

*Volume 1: *Pure Mathematics* edited by J. L. Britton, pp. 288, DFL175, \$89.50 (1992);

Volume 2: *Mechanical Intelligence* edited by D. C. Ince, pp. 226, DFL160, \$82 (1992);

Volume 3: *Morphogenesis* edited by P. T. Saunders, pp. 160, DFL170, \$106.50 (1993).

Grappling with the golem

Walter Gratzer

The Golem: What Everyone Should Know About Science. By Harry M. Collins and Trevor Pinch. Cambridge University Press: 1993. Pp. 164. £10.95, \$19.95.

WHEN James V. McConnell's theory of memory transfer bit the dust in 1970 or thereabouts, the loss was more to dinner-table discourse than to science; like the departure from the London stage of David Garrick, it "eclipsed the gaiety of nations and impoverished the public stock of harmless pleasure". For McConnell claimed to have discovered that planarian worms could be taught to turn left or right down a maze, and if they were then ground up and fed to their friends, the knowledge was absorbed, along with the meat, by the diners. So any student who was not a vegetarian could painlessly assimilate his subject by eating — raw, it is true — a portion of his professor. The conceit is pleasing enough that it deserves to be true, but science is remorseless, and eventually proved it false.

Harry Collins and Trevor Pinch, the authors of this perverse but entertaining book, are like psychiatrists — the men, you may recall, who go to the Folies Bergères to watch the audience — in that they are concerned less with what scientists do than how the scientific community as a whole reacts to an assault on its prejudices, and especially to a heterodoxy as outrageous as McConnell's. And as to that, worse was to come when George Ungar claimed to have achieved memory transfer in mammals: brain extracts from rats trained to avoid the dark by conditioning with electric shocks, he asserted, induced fear of darkness in mice. The temperacy with which workers in the field greeted the revelations of McConnell and of Ungar seems to me less startling than Collins and Pinch make out: when you have had to put up all your working life with the intellectual shackles that science forces on us all, it is galling to see a freebooting exhibitionist casually cast them aside. Collins and Pinch commend Ungar for his persistence in the face of so much hostility: when his isolation of the photophobia-inducing factor, purportedly a peptide, which he called scotophobin, from the crania of 4,000 rats failed to convince, he produced an extract from the brains of 17,000 trained goldfish, but to no greater effect. When Ungar died there was no one willing to carry the torch because, Collins and Pinch imply, the experiments were too arduous and the prospects of success too uncertain for weaker vessels to contemplate.

The argument is reminiscent of Arthur Koestler's espousal of another such hopeless cause, that of the Austrian geneticist Paul Kammerer and his midwife toads. Kammerer shot himself after being exposed as a charlatan, but Koestler preferred to believe that, unable to convince the scientific establishment, which had anathematized lamarckian inheritance, of his veracity, he committed a desperate and unpremeditated fraud. Koestler too felt that today's scientists lacked the moral fibre to take on such subtle and demanding procedures as were involved in breeding generations of fastidious animals in the laboratory, and he tried to persuade a number of zoologists to repeat Kammerer's experiments (happily without success). It was by then obvious that such an undertaking would be misconceived, for how could an acquired characteristic be imprinted into the genetic material? Collins and Pinch observe that the memory transfer experiments were overtaken by the tide of progress and by the more promising opportunities that the discovery of real brain peptides now afforded to neurobiologists.

Our authors do not go so far as to claim that McConnell, Ungar and various others of their kidney may have been right all along, but they do claim that they were rejected for the wrong reason, that is to say the unwillingness of most scientists to concede any violation of the accepted orthodoxy — the dread word 'paradigm' refuses here to be swatted. Their lopsided perspective of the scientific process leads them to take a notably benign view of the cold fusion debacle. They are kind to Stanley Pons and Martin Fleischmann, whose fault, it seems, was not reckless disregard of the accepted standards of experimental evidence — for Collins and Pinch do not allow that their claims have been proved void — but that they took on a professionally and politically entrenched élite, which would not tolerate so presumptuous a challenge to its primacy. There is also the implication, which perhaps has some substance, that Pons and Fleischmann were treated less than chivalrously by their adversaries, who, following the custom of our time, did not extend to them any sympathy when they were down and instead kicked them savagely in the slats. The plight of the fallen, moreover, was contagious, so



EVEREST from the air — the three faces of the great mountain are revealed in this photograph, taken over the summit at 35,000 feet (top, Kangshung face; right, southwest face; left, north face). The picture appears in the handsome *Everest: The Best Writing and Pictures from Seventy Years of Human Endeavour* edited by Peter Gillman (Little, Brown, \$35, £25).

when a highly regarded theoretician came up with a theoretical basis for the cold fusion phenomenon he was, it seems, denied tenure in his department. This may be so; innocent theoretical parties were certainly struck by flying fragments of custard pie in the course of a similarly farcical episode that prefigured cold fusion — the tale of polywater, another product of self-deception by reputable physicists, who should on the face of it have known better. But experimentalists might argue that the task of the theoretician, like that of the political commentator, is to predict what will happen and then to explain why the opposite occurred.

Their thesis that science is a kind of brainless, lumbering monster (the golem from Jewish mythology of the title), driven by innate ungovernable urges, leads Collins and Pinch into some strange contortions. One of their case histories, for instance, centres on Pasteur's debate with the respected naturalist, Pouchet, on the issue of whether life could be spontaneously generated, or whether, as the saw of the day had it, *omnis cellula e cellula*. It took the form of a gladiatorial confrontation before a commission of the Académie des Sciences, overwhelmingly loaded, say Collins and Pinch, against the hapless Pouchet, who eventually turned tail and ran. Further, Pasteur's experiments were flawed, most of all because the

hay-extract nutrient that he put in his sealed swan-necked flasks — still to be seen today, as void of life as a century ago — could have harboured a type of spore that is now known to resist sterilization by boiling; indeed one Henry Bastian attempted much later to rehabilitate spontaneous generation, having inadvertently discovered such spores. So it follows that Pasteur might just as well not have bothered, because seen in retrospect his experiments actually proved nothing. But this will not do: we all know that nothing retards progress so much as the right experiment at the wrong moment. What if atomic weights had been measured too accurately in the nineteenth century? Does this diminish the insights of Dalton and Prout? Or suppose that the viral polymerase that transcribes RNA into DNA had been recognized 20 years too soon. Where then would the Central Dogma be?

Collins and Pinch press the point that science is burdened by intellectual, and indeed social and political, bias (driven, for example, by its votaries' desire for funds). Scientists, if they think about such matters at all, know this well enough. Darwin, in a famous passage, remarked that he had always found it strange how anyone could doubt that observations must be for or against something if they were to be of any use. Eddington — something of a *bête noire* to Collins and Pinch — declared that one should never believe a fact until it had been confirmed by theory. But perhaps G. K. Chesterton got closest to the nub (but in a rather different context) when he wrote: "you can only find truth with logic if you have already found truth without it."

Another theme explored in this book is the hazard of pursuing very small physical effects, which lie beyond the sensitivity of existing instrumentation. Such problems attract a particular scientific type — the men with the staring eyes and the tic. Irving Langmuir in his celebrated lecture on what he called pathological science gave as one of his criteria of the genre that the reported effects were in general at the very verge of detectability and commonly irreproducible. Collins and Pinch have formulated a new law, which they call experimenter's regress: to determine whether an effect exists, a new apparatus has to be constructed, but if it reveals nothing this may signify not that there is no effect but that the apparatus is wanting, and so a vicious cycle is initiated.

This book is designed to open a window on science and its practitioners for beginners and outsiders. It gives, in my view, a distorted view of how science mostly proceeds, but for all that there are many sharp insights, the writing is deft, the stories good and there is not a boring page. The working scientist, disposed to grumble about the rigours of his vocation, can do

no better than study the absorbing account of Michelson and Morley's efforts to detect the retardation of light in the luminiferous ether by measuring its velocity along and across the Earth's axis of rotation. The first apparatus had to be moved from Berlin to Potsdam to minimize traffic vibration and even then the interference fringes could be made to vanish by the stamping of a foot 100 yards from the building. The second apparatus, built in Cleveland, Ohio, reposed in a basement and was mounted on a block of

sandstone five feet square, floating in a massive trough of mercury; set in motion by hand, it would revolve for more than an hour. But no effect was observed. The negative result acquired its retrospective éclat only after the Theory of Relativity was published and embraced by physicists. Perhaps we should not seek our rewards before the Day of Judgement. □

Walter Gratzer is in the MRC Muscle and Cell Motility Unit, Kings College London, 26–29 Drury Lane, London WC2B 5RL, UK.

Flying squid and spinning seeds

R. McNeill Alexander

The Biokinetics of Flying and Swimming. By Akira Azuma. Springer: 1992. Pp. 265. DM280, £150, \$198.

SQUID swim fast by jet propulsion, using enhanced breathing movements to squirt water out of their gill cavity so as to drive them forwards or back (backwards is faster). That is how ordinary squid travel, but the flying squid of the Indian Ocean also leaps into the air to glide like a flying fish. Most people (including professional zoologists) will be puzzled to identify the



This ladybird, *Coccinella 7-punctata*, wings its way into John Brackenbury's *Insects in Flight* (Cassell), which was shortlisted for this year's Rhône-Poulenc Science Book Prize.

subject of the photograph on the cover of Akira Azuma's book — a shoal of squid in flight — and all will be struck by the oddity of these animals. Flying fish have enlarged fins placed like the wings of an aeroplane, but the flying squid has its aerofoils at the extreme ends of its cigar-shaped body, a pair of fins at the anatomical rear end (which travels in front) and membranes stretched between the tentacles of the anterior end (which follows behind).

Azuma reviews the swimming and flight of a remarkable range of organisms, several of them as odd as the flying squid. He discusses the flight of birds, bats, many different kinds of insect, flying squirrels, sycamore seeds and spiders on gossamer,

as well as squid. His chapters on swimming deal with the whole range of animal sizes from microorganisms to whales, all shapes of fish, and wind-surfing scallops. But he leaves one with the impression that he is interested in organisms only as problems in theoretical mechanics, that he finds the animals less interesting than the equations.

The initial impression given by the wonderful cover picture is quickly modified when the book is opened: the first ten pages, in double columns, are filled by a list of algebraic symbols. Most pages have a few equations, which are generally neither complicated nor difficult, though readers who do not already know (for example) what a Kármán vortex is had better keep a hydrodynamics text handy.

The strength of this book is its richness as a source of data and equations. Many graphs present potentially useful data that are hard to find elsewhere, such as force coefficients for rectangular and triangular aerofoils of low aspect ratio. And tables present the basic equations for the analysis of (for example) the helical beat of a flagellum, while others give data on the dimensions and performance of (again, one of many examples) spinning seeds.

The weakness of the book is its lack of reference to experimental physiology. There are diagrams of theoretically expected vortex patterns in the wake of flying birds but no illustrations of the actual patterns revealed by experiments with helium-filled soap bubbles. There is a table showing that squid that swim slower than similar-sized fish nevertheless use more energy, but no indication of how this information was obtained.

That said, those of us who research and teach in the field of animal locomotion will refer frequently to this book, and teachers of applied mathematics looking for unusual problems will find it a rich source of ideas. □

R. McNeill Alexander is in the Department of Pure and Applied Zoology, University of Leeds, Leeds LS2 9JT, UK.

Clues to the laws of ecosystems

David J. Beerling

Climate Change and Its Biological Consequences. By David M. Gates. *Sinauer*: 1993. Pp. 280. \$18.95, £16.95 (pbk).
Biotic Interactions and Global Changes. Edited by Peter M. Kareiva, Joel G. Kingsolver and Raymond B. Huey. *Sinauer*: 1993. Pp. 480. \$65, £58.95 (hbk); \$32.95, £29.95 (pbk).

ARE there fundamental laws governing the behaviour of ecosystems? If so, how long will it be before we can expect to discover them? Unfortunately, unlike physicists and chemists, ecologists have few, if any, general laws, and none that can be applied to the problem of predicting the responses of ecosystems to global climate change.

Ecologists are now trying to catch up with the physical sciences. The perceived threat of global change is forcing them to focus, in unprecedented detail, on how ecosystems function and interact with the environment. Simple correlations are just not sufficient — global change has redirected ecologists' attention towards finding generalities that are both precise and realistic.

D. M. Gates provides a readable and up-to-date summary of many of the aspects of global ecology required for an understanding of the present state of play. The coverage of topics is necessarily diverse, and brief, ranging from the astronomical theory of orbital forcing as a mechanism for Quaternary climatic change, to plant physiology, drought and El Niño events. Emphasis is placed on vegetation responses to global change; past vegetational response, the use of forest models and ecosystem responses comprise most of the book. Each topic is given a competent treatment that will be of use to students new to the subject but adds little to the knowledge of those already working in the field.

P. M. Kareiva and colleagues have edited a very different volume, pitched more at researchers. Concentrating only on ecosystem responses to global changes (excluding oceanic interactions and biogeochemical cycles), they have compiled chapters written by experts in their fields that attempt to develop new concepts rather than simply review the existing literature. The volume arose from a workshop on the effects of global change on populations and communities of plants and animals. In general, the approach works well. Of particular interest is the chapter by David Tilman on plant responses to carbon dioxide and that by Stephen Pacala and George

Hurt on the difficulties of forest models. Other chapters concentrate on stream communities, birds and interactions between plants and animals. An evolutionary section deals with species extinction caused by variation in the rates of climate change, migration and habitat fragmentation. Moreover, it tackles the more subtle concept of microevolution and the genetic variation of populations. My only criticism of the book as a whole is the parochial selection of authors: 40 out of the 45 are based in the United States.

Biotic Interactions and Global Changes emphasizes the search for generalities and provides several clear pointers and clues to where to go from here. The prospect of discovering the laws of ecology must now seem brighter thanks, ironically, to the prospect of a warmer world. □

David J. Beerling is in the Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2UQ, UK.

Gravity waver

Eric G. Adelberger

Gravitational Experiments in the Laboratory. By Y. T. Chen and Alan Cook. *Cambridge University Press*: 1993. Pp. 268. £40, \$59.95.

THE past decade has seen a small renaissance in laboratory tests of gravity, stimulated by a new way of interpreting the remarkably sensitive results of these ingenious mechanical experiments. The experiments are now viewed not so much as testing the theory of gravity (general relativity has passed all tests with flying colours, including those derived from the beautifully precise binary pulsar data) but as searches for possible new and very feeble interactions that may lie hidden 'underneath' gravity. There is a need for a new book aimed at those who seek an overview of the field or wish to understand the power and limitations of the techniques. *Gravitational Experiments in the Laboratory* will be useful to such readers, but is oddly flawed as well.

With commendable generality, Chen and Cook begin by discussing a linear oscillator in contact with a thermal bath, and then consider other possible disturbances of the oscillator. They apply this foundation to tests of the universality of free fall (the weak equivalence principle), to tests of the $1/r^2$ law (the gravitational Gauss law) and finally to the notoriously difficult measurements of the newtonian gravitational constant G . (Although Boys claimed in 1889 that his torsion balance would be able to determine G to a precision of one part in 10^4 , this goal has not yet been attained.)

The authors give a good account of the history of the field. But it has a curiously 'old-fashioned' flavour because it omits most of the developments since the reanalysis by E. Fischbach and colleagues in 1986 of the classic Eötvös equivalence principle data, which seemed to provide evidence for a new fundamental interaction or 'fifth force'. The book also contains serious errors. For example, tests of the weak equivalence principle that compare the gravitational accelerations of different materials towards the Sun are explained as relying on tidal forces. This is wrong: these experiments do not probe tidal forces, which vary as $1/r^3$, but the main gravitational force, which varies as $1/r^2$. Another serious error concerns F. Witteborn and W. Fairbank's test of gravity using freely falling electrons. Their observation that electrons in the centre of a hollow conducting tube do not fall has never been confirmed. And L. Schiff and M. Barnhill's theoretical explanation for this 'fact', which Chen and Cook use as evidence that electrons feel the expected gravitational forces, is now thought to be in error.

There are also strange quirks in the discussion of experimental matters: circuit diagrams that make no sense and subsequent discussions that are therefore impenetrable; quantities given with many significant figures for no purpose; order-of-magnitude errors (a capacity of 0.5 picofarads is equated to that of 1 metre of coaxial cable); and crucial points misstated (in Mössbauer tests of the gravitational redshift, the temperatures of source and absorber have to be carefully controlled not because the first-order Doppler effect broadens the line but because the second-order effect shifts it).

A careful reader able to separate the wheat from the chaff will find much useful material in this book. However, a student wishing to learn about the field will need guidance on what not to accept, while someone wanting to know the status of the field will have to look elsewhere to learn about recent developments. □

Eric G. Adelberger is in the Department of Physics, University of Washington, Seattle, Washington 98195, USA.

■ Cambridge University Press has also just published a new edition of *Theory and Experiment in Gravitational Physics* by Clifford M. Will, price £19.95, \$37.95. For a review of the first edition, see *Nature* **299**, 284; 1982.

Correction

In his review of *DNA on Trial* (*Nature* **363**, 122; 13 May 1993), John F. Y. Brookfield incorrectly attributes a chapter written by P. Rabinow to the editor of the volume, P. R. Billings. Our apologies for the error.

Interstellar grains in meteorites

Ulrich Ott

Primitive meteorites contain grains of silicon carbide, graphite and diamond formed outside the Solar System and probably before its birth. The isotopic compositions of these grains provide a record of stellar nucleosynthesis and of condensation processes near carbon stars; the fact of their survival places constraints on conditions in the solar nebula and early Solar System. The search is now on for other surviving stellar condensates, such as nitrides and oxides.

THE chemical elements and their isotopes in the Solar System largely originate from nucleosynthesis in stars¹⁻³. Since the Solar System's formation (and before), the isotopic abundance ratios established by these nucleosynthetic processes have been modified by several well understood secondary processes: radioactive decay of long-lived radionuclides (such as ⁴⁰K, ²³⁸U and ²³⁵U, ⁸⁷Rb), interaction with cosmic rays, and mass-dependent fractionation effects that occur in physical or chemical processes such as diffusion, phase changes, thermodynamic equilibration or chemical kinetics. (Mass-independent fractionation of isotopes can also occur under certain conditions⁴⁻⁶, but is not important in the current context.) Apart from these effects, it has traditionally been assumed — and in the 1960s and early 1970s was regarded by many as “one of the few assumptions that can be considered well-justified and firmly established”⁷ — that the Solar System is isotopically homogeneous, as was its predecessor, the solar nebula. The planets and other objects in the Solar System were thought to have formed from a “well-mixed primordial nebula of chemically and isotopically uniform composition”⁷, with temperatures in this nebula so high that no solids survived.

But even then, evidence to the contrary was already available, in the shape of anomalous abundances, not explained by the above processes, of certain isotopes of the noble gases xenon and neon. The anomalous component now known as xenon-HL, had been discovered in the carbonaceous chondrite Renazzo⁸ in 1964 (although at that time it was ascribed to one of the ‘well-understood’ processes, fission) and in 1969 ‘neon-E’ was found in several primitive meteorites⁹. Both of these isotope abundance anomalies, not taken very seriously outside the noble gas community at the time, are now known to be carried by interstellar grains of diamond, graphite and silicon carbide — stardust that formed around stars in their late stages. Although originally identified because of isotope abundance anomalies in the noble gases, these phases have since been shown to be

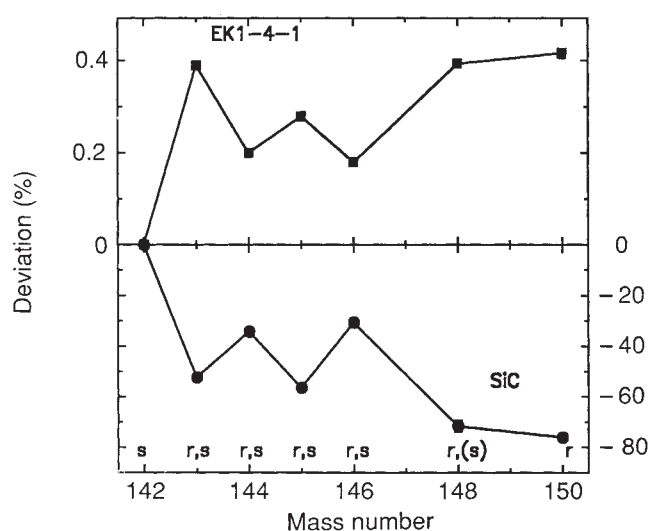


FIG. 1 Nd isotope abundance anomalies (deviations of isotopic ratios from normal; normalized to ¹⁴²Nd) in interstellar silicon carbide⁵⁶ (bottom) compared with that in FUN inclusion EK1-4-1 (ref. 102; top). The anomalies are of opposite sign, and that for interstellar SiC is two orders of magnitude larger. In the case of such large anomalies, corrections for mass-dependent fractionation, for example instrumental mass discrimination, are not important.

isotopically anomalous also in their major elements — that is, in the very stuff of which they are made.

It was not these anomalies, but a string of discoveries of isotope abundance anomalies in other elements, starting with the ¹⁶O anomaly in 1973 (ref. 10) and most prominent in the high-temperature Ca–Al-rich inclusions (CAIs) of the primitive Allende meteorite (notably the ‘FUN’ inclusions¹¹, named for fractionation plus unknown nuclear effects), that led meteoritists to accept that the solar nebula could not have been perfectly homogeneous. There is an important difference, however, between the isotope abundance anomalies found in the high-temperature inclusions in primitive meteorites and those observed in interstellar graphite, diamond and silicon carbide. Whereas the former, in all likelihood, constitute material that was processed or reprocessed within the Solar System or solar nebula, and which formed from reservoirs that were not in full equilibrium with the rest, the latter clearly constitute what we call pre-solar material. They formed outside the Solar System and were incorporated into it with the memory of their nuclear sources intact (for major as well as trace elements). As a consequence, the size of the anomalies is very different: whereas in the case of the CAIs deviations from ‘normal’ isotopic composition are commonly measured in per mil or ε-units (parts in 10³ or 10⁴), the anomalies in interstellar grains are usually

TABLE 1 Properties of surviving interstellar grains in meteorites

Phase	Abundance (p.p.m.)*	typical size (μm)	Associated Ne, Xe	Other anomalous elements†
Graphite	<1	~1–6	Ne-E(L)	C,N,O,Mg, Si(?)
Diamond	~400	~0.002	Ne-A2, Xe-HL	N,Ba(?), Sr(?)
SiC	~5	~0.2–10	Ne-E(H), Xe-S	C,N,Mg,Si,Ca,Ti, Sr,Ba,Nd,Sm

Ne-E(L) and Ne-E(H) are two types of Ne-E (essentially pure ²²Ne), originally designated L for low density, low release temperature, and H for high density, high release temperature, respectively. Ne-A2 is the typical ‘planetary’ trapped component found in meteorites. Xe-HL is characterized by excesses in the light and heavy Xe isotopes relative to those of intermediate mass. The isotope pattern of Xe-S corresponds to what is expected to be produced in the slow neutron capture process (s-process) of nucleosynthesis.

*By mass; typical value for C2M meteorites (Murchison).

†In addition to the noble gases.

orders of magnitude larger (see, for example, the neodymium anomalies in Fig. 1).

A few of the characteristics of interstellar diamond, graphite and silicon carbide are listed in Table 1. Here I will focus mainly on what the isotopic signatures reveal about the nucleosynthetic sources of the elements in the grains. Stars on the asymptotic giant branch in the Hertzsprung–Russell diagram (AGB stars; Fig. 2) have probably provided most of the silicon carbide found in primitive meteorites. Moreover, AGB stars in their thermally pulsing phase (TP–AGB stars), alternately burning hydrogen and helium in shells (see, for example, refs 12, 13), are thought likely to be a chief source of the elements produced in the slow neutron capture process (s-process) of nucleosynthesis; thus, the isotopic compositions of the heavy trace elements in silicon carbide provide a wealth of information, otherwise unobtainable in such detail and with such accuracy, regarding this process. Supernovae may be required to explain the isotopic features of the most abundant of the phases, diamond, as well as some exotic silicon carbide grains.

The fact that (at least) these three interstellar phases have survived has implications for the history of the early Solar System, the potential of which we are just beginning to explore.

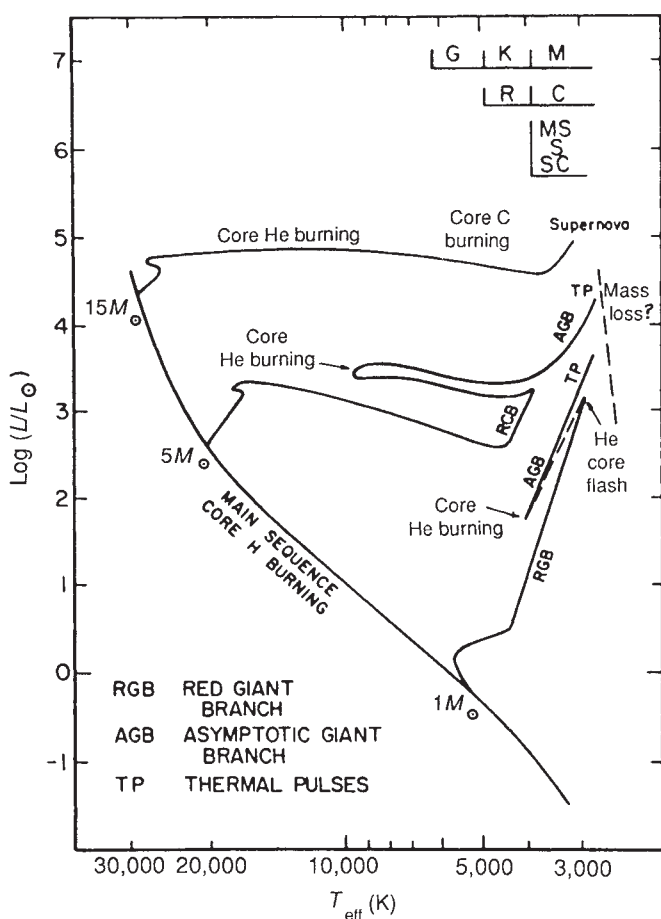


FIG. 2 Hertzsprung–Russell (effective temperature versus luminosity) diagram showing the development of stars with 1 and 5 solar masses into the AGB star region. After spending most of their life on the main sequence, burning hydrogen to helium, most stars (all less massive than about 8 solar masses) contract their cores and expand their envelopes, moving into the red giant region (red giant branch), from which, after the onset of helium burning (helium flash) they jump back, evolving into the red giant region a second time along a similar track (asymptotic giant branch). At the end of their lifetime these stars develop strong winds and finally eject their envelopes as planetary nebulae, ending their lives as white dwarfs. Figure courtesy of V. V. Smith.

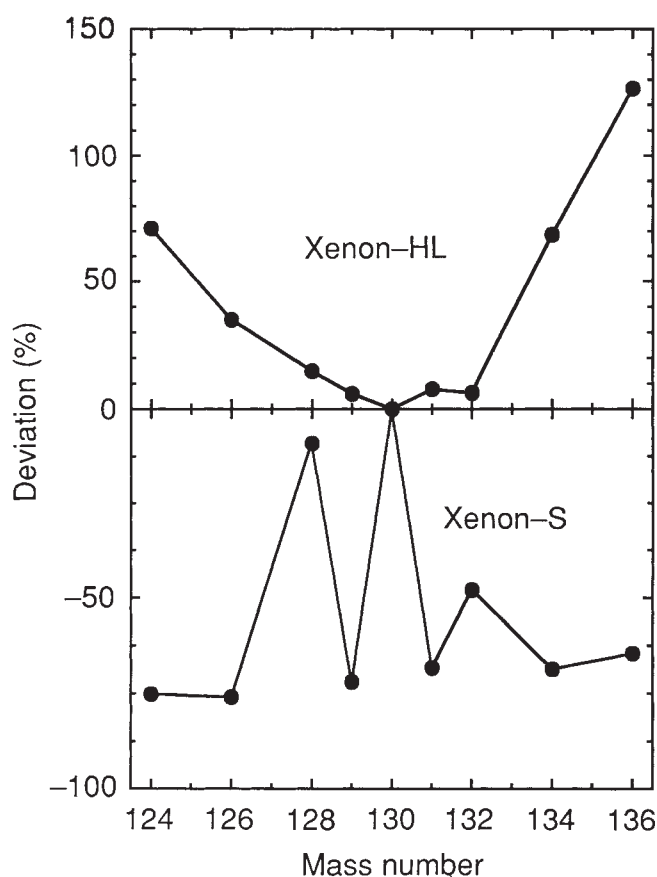


FIG. 3 Abundance pattern of xenon-HL (ref. 83) and s-process xenon⁶¹. Shown here are the deviations in per cent of the ¹³⁰Xe-normalized ratios from those in solar-wind xenon as measured in lunar soil.

Other types of stellar condensate may also have survived in meteorites — indeed, two grains of interstellar corundum (aluminium oxide) have recently been found^{100,101}. The field is currently in a rather explosive phase of discovery, in which often new findings seem to add complexity, rather than yield new insight; nevertheless, there are certain first-order conclusions that can be drawn concerning the sources of the interstellar grains and their incorporation into Solar System material. For more details, and for aspects of the story not covered here, see refs 14–17.

The quest for noble-gas host phases

Noble gases have played an important role in establishing the interstellar origin of diamond, graphite and silicon carbide. All three phases are 'tagged' by noble gases that are isotopically unusual. Almost pure ²²Ne (Ne-E) is carried by graphite (Ne-E(L)) and silicon carbide (Ne-E(H)); diamond carries xenon-HL, which shows excesses in the light and heavy isotopes relative to those of intermediate mass; and both silicon carbide and graphite contain xenon (xenon-S) of the composition expected to be produced by the s-process (Fig. 3).

Because meteorites contain only minute quantities of noble gases (relative to the rock-forming elements, xenon is only $\sim 10^{-4}$ and helium $\sim 10^{-9}$ of its solar abundance), the isotopic compositions of noble gases carried by interstellar grains shine through even when the grains are 'diluted' with extraneous (largely noble-gas-free) matter. It was the quest for the phases that carried the observed noble gas isotope anomalies (primarily by workers at the University of Chicago) that finally led to their identification. The breakthrough came in 1975, when it was noticed that in primitive meteorites most of the trapped noble gases were contained in a small, HF/HCl-resistant fraction¹⁸ and

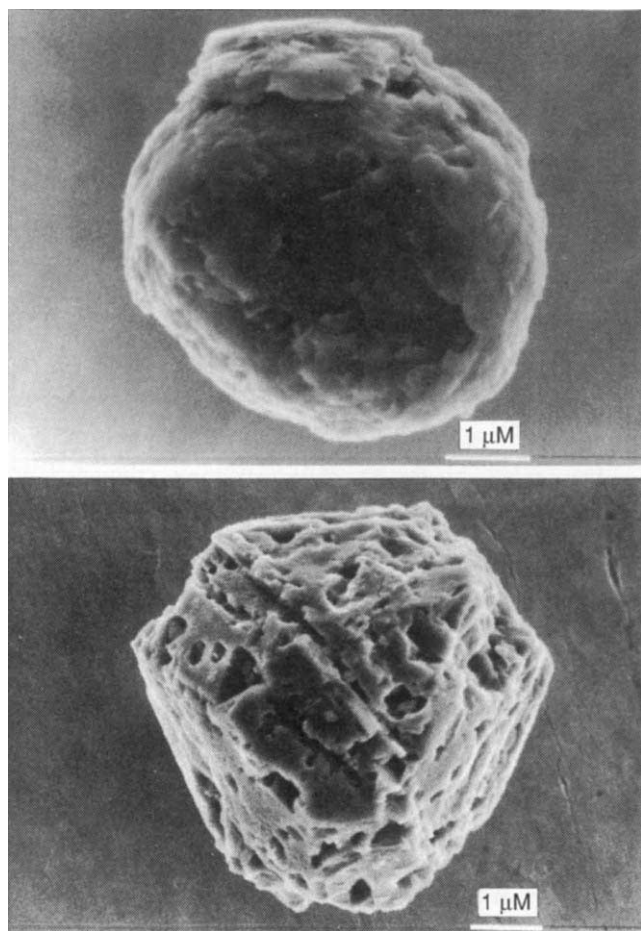


FIG. 4 Scanning electron micrographs of grains of interstellar graphite (top) and silicon carbide (bottom). Courtesy of S. Amari.

that appropriate further treatment could largely remove the ordinary component and thereby enrich anomalous components. This finally led to the identification in 1987 of tiny (<3 nm diameter) diamonds as the carrier phase of xenon-HL¹⁹, and, as a bonus, the discovery of interstellar silicon carbide^{20,21} and graphite²².

Physical properties and astronomical observations

Diamond, graphite and silicon carbide are refractory (temperature-resistant) carbon compounds which could have formed only under reducing conditions. This limits the possible stellar sources to carbon-rich red giant stars (thought to be the main source of interstellar carbon-rich dust²³), novae, carbon-rich shells from supernovae and massive, mass-losing, carbon-rich, Wolf-Rayet stars (WC stars)²³. Graphite, amorphous carbon and polycyclic aromatic hydrocarbons (PAHs) have long been thought to be present in interstellar dust because of spectral absorption features²³, with an emission feature at ~ 11.3 μm indicating the presence of SiC in the atmospheres of carbon stars²³. As shown below, carbon-rich red giants during the AGB phase could account for many of the nucleosynthetic features of the 'typical' SiC grains found in meteorites, so they may have supplied most of these grains. Finally, diamond-like material has been tentatively identified in dense molecular clouds surrounding currently forming stars^{24,25}. Although technically possible, formation of diamond by shock pressure seems a less likely source of the meteoritic pre-solar diamond grains than direct formation by a process analogous to chemical vapour deposition¹⁴. Another possibility may be annealing of small hydrocarbon grains to produce tiny diamonds, caused by

absorption of far-ultraviolet photons near a supernova²⁶.

There seem to be differences between the size distribution of interstellar grains in meteorites^{14,27} and those inferred from astronomical observations²⁸, which may reflect the different nature of the grain populations²⁷ and also may be telling us something about grain formation in stellar atmospheres. Astronomical observations of interstellar grains indicate power-law distributions with slope 3.3–3.6, extending from 0.005 μm to ~ 1 μm for graphite and from 0.025 to 0.25 μm for other phases. Meteoritic diamond grains seem to be at the lower end, SiC and graphite (Fig. 4) at the upper end of these size ranges (Table 1), with distributions that are of log-normal rather than power-law type. For the larger silicon carbide grains the size distribution is also compatible with a power law, but in this case the slope of ~ 5.7 would be significantly different from the one inferred from the astronomical observations²⁷.

Graphite and silicon carbide

Consideration of the isotopic characteristics of graphite and silicon carbide — particularly for carbon and nitrogen (which occur at about the per cent level in both) and magnesium — requires a similar approach, so it is convenient to discuss them together. Because there is much more complete information for SiC, however, discussion of this phase will dominate in what follows.

The carbon, silicon and titanium isotopic compositions of most individually analysed SiC grains define a 'normal' range for these anomalous grains (albeit with some clustering^{29,30}), but some rare, 'exotic' subtypes stand out (grains 'X' and 'Y'; see below). In contrast to the light elements, which exhibit complex patterns, the isotopic compositions of heavy trace elements in both graphite and SiC are easier to interpret and the results, in particular for SiC, can provide information on details of the s-process of nucleosynthesis. Overall, the situation seems clearer for silicon carbide than for graphite, and there are a number of indications that most of the grains may have originated in stars of the asymptotic giant branch (AGB stars). Perhaps the strongest argument is that astronomical observations of AGB stars indicate elemental and isotopic abundance patterns of the CNO elements and enhancement of s-process nuclides (see refs 12, 31, 32) similar to those observed in the meteoritic grains. Nevertheless, the situation is not entirely straightforward and certain problems remain: for example, models based on adopted neutron density profiles^{32–34}, although very successful in reproducing those isotopic signatures that are due to the action of the s-process (see below), are still just parametric studies and it must be borne in mind that, with one exception, the most detailed stellar evolution calculations for AGB stars have failed to predict neutron density profiles that would produce a significant s-process (see ref. 35 and, for a critical appraisal, refs 36, 37). And even the parametric calculations so far fail to explain several of the meteoritic observations, notably the silicon and titanium isotopic compositions.

Light ($A < 60$) elements. The first carbon isotope measurements on samples enriched in what was later identified as SiC were done by conventional mass spectrometry and indicated an unusual isotopic composition with $\delta^{13}\text{C} \approx +1,400\text{‰}$ (that is, $^{12}\text{C}/^{13}\text{C} \approx 40$ rather than the normal terrestrial value of ~ 89)³⁸. Ion microprobe work on pure SiC samples confirms this value as an average³⁹, but also shows both $^{12}\text{C}/^{13}\text{C}$ and $^{14}\text{N}/^{15}\text{N}$ to be highly variable from grain to grain^{30,39} (Fig. 5a). Excluding the exotic grains X and Y, the $^{12}\text{C}/^{13}\text{C}$ ratio ranges from ~ 2 to ~ 200 (0.02 to 2 times the solar value of 89), with most grains plotting in the range 60 ± 15 . Nitrogen is equally variable, ranging from ~ 80 to $\sim 5,000$ (0.3 to 20 times the solar value of 272). There is some clustering in carbon composition which suggests the existence of subgroups, but each group would still have large variations (more than an order of magnitude) in $^{14}\text{N}/^{15}\text{N}$ (Fig. 5a). The evidence for clusters is especially strong for the very

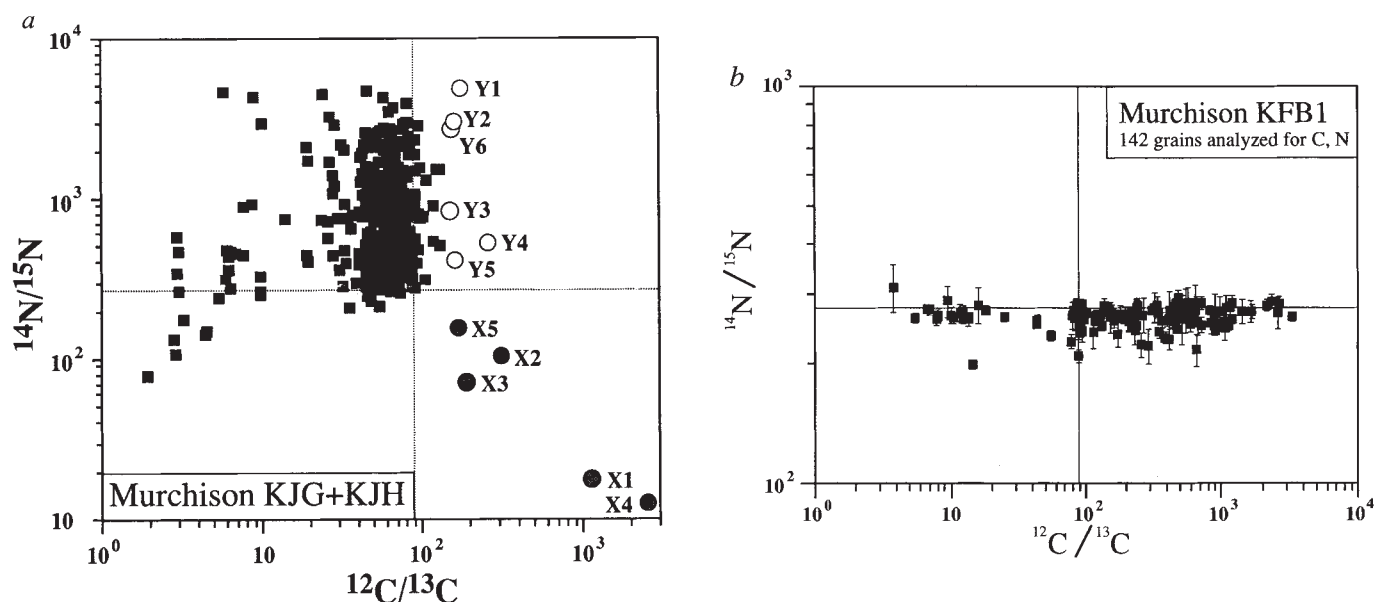


FIG. 5 $^{14}\text{N}/^{15}\text{N}$ plotted against $^{12}\text{C}/^{13}\text{C}$ measured in individual grains of interstellar silicon carbide (a) and graphite (b). Figures courtesy E. Zinner.

large (up to $\sim 20\ \mu\text{m}$) grains from one of the earlier acid-resistant residues (the 'L series') where three clusters of carbon and silicon isotopic composition (of which one is isotopically normal) clearly stand out^{29,30}. The later and much more intensively studied 'K series', however, shows little evidence for abundant presence of such large, clustered, grains³⁰, and we may not have 'a completely representative sample' of Murchison SiC yet⁴⁰.

The situation for graphite²² is somewhat different (Fig 5b shows an example; note the difference in scales). $^{14}\text{N}/^{15}\text{N}$ is, on the average, lower than but close to the terrestrial value, and deviations in either direction do not exceed a factor of 3 (the value inferred from analyses of ensembles of many grains by conventional mass spectrometry^{16,41} is $\delta^{15}\text{N} > 252\text{‰}$, or $^{14}\text{N}/^{15}\text{N} > 220$). Variation of $^{12}\text{C}/^{13}\text{C}$, on the other hand, is much larger than in the 'normal' SiC grains, extending from ~ 3 to $\sim 4,000$, or 0.03 to > 20 times the terrestrial value.

Nucleosynthesis processes that produce carbon and nitrogen isotopic compositions qualitatively similar to these are^{15,30,39}, hydrostatic hydrogen burning via the CNO cycle, as it occurs in red giant stars, for low $^{12}\text{C}/^{13}\text{C}$, high $^{14}\text{N}/^{15}\text{N}$ in the upper left (northwest) quadrant; explosive hydrogen burning (possible sites: novae) for low $^{12}\text{C}/^{13}\text{C}$ coupled with low $^{14}\text{N}/^{15}\text{N}$ (lower left, southwest); and helium burning that takes place in massive stars for high $^{12}\text{C}/^{13}\text{C}$ in the right-hand (eastern) sectors.

Obviously, it is not possible to identify a single process that can reproduce the isotopic signatures in these interstellar grains, especially in the case of graphite with its large variations of $^{12}\text{C}/^{13}\text{C}$ coupled with modest variations in $^{14}\text{N}/^{15}\text{N}$. All the above processes may have affected the observed compositions through contributions from different zones or stages in a star's evolution, formation in different stars, or differences in the initial composition that were not fully erased in the subsequent evolution.

The case of SiC seems simpler than that of graphite, as most grains occupy a position in the northwest quadrant indicating significant contributions from the CNO cycle. But only a small fraction of the grains approaches the equilibrium value for this process, $^{12}\text{C}/^{13}\text{C} \approx 3.4$ (ref. 42). Oddly enough, the low- $^{12}\text{C}/^{13}\text{C}$ data points do not show $^{14}\text{N}/^{15}\text{N}$ values approaching the CNO cycle equilibrium value, as might be expected if contributions from this process also dominated the nitrogen isotopic ratios. This value is in the range $(2-5) \times 10^4$, depending on the exact temperature at which the process operates ($^{12}\text{C}/^{13}\text{C}$ shows virtually no temperature dependence)⁴³. In fact, the low- $^{12}\text{C}/^{13}\text{C}$ grains have, in general, lower $^{14}\text{N}/^{15}\text{N}$ than those in the region around 70 where most grains plot (Fig. 5a).

About a quarter of the graphite and more than half of the SiC grains analysed show large excesses of ^{26}Mg . The most likely source is decay of radioactive ^{26}Al ($T_{1/2} = 0.7$ Myr) within the grains^{39,44}. Inferred $^{26}\text{Al}/^{27}\text{Al}$ ratios at the time of grain formation range up to 10^{-2} (excluding the extraordinary grains) for SiC and up to 10^{-1} for graphite, with no simple relationship to any of the other measured isotope ratios. Possible nucleosynthetic sources for this putative extinct ^{26}Al include hydrogen burning in AGB or Wolf-Rayet stars, explosive hydrogen burning (novae) or explosive carbon or neon burning (supernovae) although it seems difficult to produce the highest values in the latter⁴⁴.

Silicon in SiC shows a more restricted range of isotopic compositions^{29,30,39}. Most grains fall in the range -60 to $+200\text{‰}$ for $\delta(^{29}\text{Si}/^{28}\text{Si})$ and from -40 to $+130\text{‰}$ for $\delta(^{30}\text{Si}/^{28}\text{Si})$. The values lie close to a line with slope ~ 1.3 in a plot of $\delta^{29}\text{Si}$ against $\delta^{30}\text{Si}$ (Fig. 6; see legend of this Figure for information on δ notation). Note that this line does not pass through the origin. The s-process has primarily been considered as the process that may alter the silicon isotopic composition (enriching ^{29}Si and ^{30}Si relative to ^{28}Si)³³, but it has difficulties in reproducing the slope of ~ 1.3 (see also below).

The isotopic composition of Ti, expressed as deviations from the normal composition, shows a V-shaped pattern with both the light ($^{46,47}\text{Ti}$) and heavy ($^{49,50}\text{Ti}$) isotopes enriched relative to the intermediate (most abundant) isotope ^{48}Ti (refs 30, 45). Qualitatively, this corresponds to the pattern that would be produced by the s-process, although, again, the agreement is not perfect.

In graphite, oxygen of anomalous isotopic composition has been found⁴⁶. About 1/4 of all grains analysed for oxygen show large excesses of ^{18}O , with $^{18}\text{O}/^{16}\text{O}$ enhanced up to an order of magnitude above the terrestrial value of 2×10^{-3} ($^{17}\text{O}/^{16}\text{O}$ seems normal within the rather large errors of, typically, some tens of per cent). This virtually monoisotopic excess of ^{18}O probably arises from the (α, γ) reaction on ^{14}N , which is abundantly present after operation of the CNO cycle⁴⁶. The reaction product ^{18}F decays into ^{18}O . Because in the same reaction chain ^{22}Ne is produced through a similar (α, γ) reaction, it might be interesting to explore a possible correlation between the abundances of the two nuclides. It is suggestive that the fraction of graphite grains carrying Ne-E(L) also seems to be of the order of $\sim 1/4$ (ref. 47).

Anomalous anomalies. Of the many SiC particles from the Murchison meteorite that have been individually analysed, a

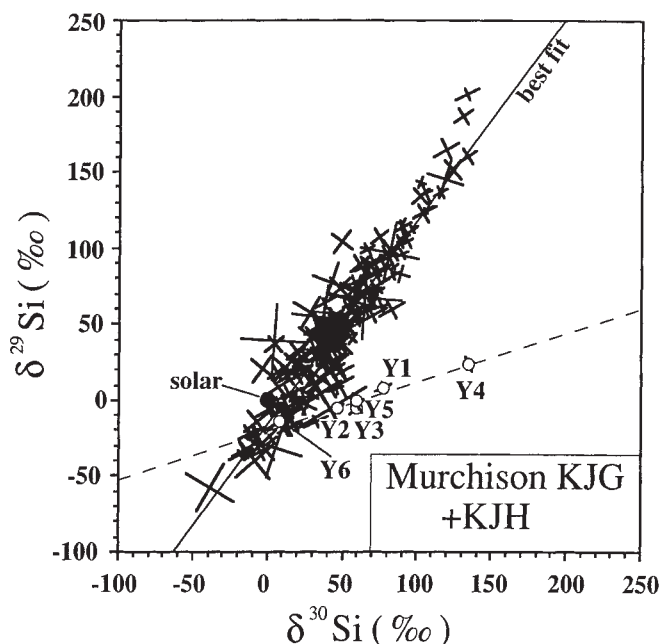


FIG. 6 Silicon isotopes measured in individual silicon carbide grains. Plotted are the deviations in per mil (‰) from the terrestrial values: $\delta(^{29}\text{Si}/^{28}\text{Si})$ on the ordinate against $\delta(^{30}\text{Si}/^{28}\text{Si})$ on the abscissa. The data scatter around a line with a slope ~ 1.3 , which does not pass through the origin. $\delta(^{29}\text{Si}/^{28}\text{Si}) = [(^{29}\text{Si}/^{28}\text{Si})_{\text{sample}} / (^{29}\text{Si}/^{28}\text{Si})_{\text{terr}}] - 1$. Figure courtesy of E. Zinner.

few stand out. They have been labelled grains X (refs 48, 49) and grains Y (ref. 49) (five and six grains, respectively, out of 720 grains analysed in ref. 30; see also Figs 5a, 6). Both types are characterized by high $^{12}\text{C}/^{13}\text{C}$, whereas in most such grains this ratio is low. In type X, such carbon is accompanied by ^{15}N -rich nitrogen, putting these grains into the southeast quadrant of Fig. 5a, whereas the grains Y plot in the northeast quadrant ($^{14}\text{N}/^{15}\text{N}$ up to $\sim 5,000$). Silicon isotopes for the grains X show large negative δ values down to $\delta(^{30}\text{Si}/^{28}\text{Si}) \approx -600$ ‰. Four of the five grains X define a line passing close to the origin and having a slope ~ 0.6 .

The grains Y plot on a line with a similar slope (~ 0.4 ; Fig. 6), but with an offset to the right and down from the main trend, with a much smaller range and closer to the origin. The slope is close to what is predicted from models of s-processing in TP-AGB stars of low mass³³. Their high $^{12}\text{C}/^{13}\text{C}$ need not be an obstacle to an AGB star origin, as this could be accommodated by assuming contributions from the essentially pure ^{12}C of the underlying helium-burning shell.

The problem of the grains X is more complex, possibly because more data are available. None of the stellar sources where the chemical environment would allow the condensation of SiC can provide a simple explanation for the observed isotopic compositions. Supernovae have nonetheless been regarded as the most likely candidates⁴⁸. The apparent ability to account for the effects in calcium and titanium may be the most attractive feature of this suggestion: strong enhancements up to $\sim 1,000$ ‰ in ^{49}Ti relative to ^{48}Ti accompanied by smaller excesses in ^{50}Ti have been observed in all grains X analysed for titanium, and calcium with essentially normal^{42,43} $\text{Ca}/^{40}\text{Ca}$ but a huge over-abundance of ^{44}Ca ($\delta(^{44}\text{Ca}/^{40}\text{Ca}) \sim +3,000$ ‰) is present in grain X2 (ref. 48). *In situ* decay of radioactive precursors ^{44}Ti ($T_{1/2} \approx 47$ years) and ^{49}V ($T_{1/2} \approx 331$ days) produced by explosive nucleosynthesis⁴⁸ may be a reasonable explanation for the effects seen in ^{44}Ca and ^{49}Ti . The main snag about a supernova origin is that it does not account for the very high $^{26}\text{Al}/^{27}\text{Al}$ ratios (at least for SN type II, these are up to 2 orders of magnitude higher than expected), unless the various shells are mixed in the right proportions.

Interestingly, the elemental ratios V/Ti and Ti/Ca are not extreme and thus not really favourable for finding anomalies of the extinct radioactivity type, which, in general, are observed only in phases where the ratio of parent to daughter element is high. Nevertheless the small degree of fractionation may be useful, as it gives lower limits on the production ratio of unstable ^{44}Ti and ^{49}V relative to the stable Ti and V isotopes, which in turn could yield clues to the nucleosynthesis process involved.

Silicon carbide and the s-process. In contrast to the complex situation for the light elements, the observed isotopic compositions of heavy trace elements found in graphite and SiC can be interpreted as simple mixtures of two components — one near 'normal', and one that originates from the s-process. The latter (for isotopes that are produced by the s-process only) accounts for about two-thirds of the purest SiC samples (compare with Fig. 1). But except for the noble gases (see ref. 50), measurements have been made on SiC only, and because of the low abundance of these elements, only on larger samples (aggregates) and not single grains, so their composition in the exotic grains is unknown. Elements analysed in SiC include (besides the noble gases) strontium^{51,52}, barium^{53,54}, and neodymium/samarium^{55,56}.

From the assumption of a simple two-component mixture, we can deduce the isotopic compositions of the pure s-process component. The isotopic compositions of s-process krypton and barium, in particular, put constraints on details of the process. For example, because of the extremely small neutron capture cross-section of ^{138}Ba , the ratio $(^{138}\text{Ba}/^{136}\text{Ba})_s$ in the s-process component is a sensitive measure of the neutron dose (neutron exposure) experienced by the seed nuclei in the process. It indicates an exposure which is within the range derived from observations of red giant stars^{31,32}, but almost a factor of two lower ($0.14\text{--}0.17$ rather than ~ 0.30 millibarn $^{-1}$, for $kT = 30$ keV) than inferred for the 'main component'⁵⁷ from the solar abundances of the nuclides $A > 90$ produced in the s-process. This fact indicates that this is not just average Solar System s-process material.

S-process krypton varies in composition as a function of krypton release temperature during combustion of SiC⁵⁸ and as a function of grain size⁵⁹, in the abundances of the isotopes ^{80}Kr and ^{86}Kr . The production of these isotopes is via unstable ^{79}Se and ^{85}Kr , respectively, the half-lives of which, under stellar conditions, are comparable to the timescale relevant for the s-process. Hence both neutron capture and β -decay of these isotopes do occur at non-negligible rates ('branching'), and their abundance, relative to the other krypton isotopes, depends on the exact conditions such as neutron density and, in the case of ^{80}Kr , on temperature (because of the temperature dependency of the ^{79}Se half-life). The observed variations show that not all s-process krypton in meteoritic silicon carbide was produced under identical conditions. The generally low $^{80}\text{Kr}/^{82}\text{Kr}$ ratio indicates less β -decay of ^{79}Se (the branch that leads to production of ^{80}Kr) than was thought typical of the s-process, and therefore lower temperature^{53,58}.

These observations confirmed doubts about the role of intermediate-mass AGB stars as the prime source of the s-process nuclides ascribed to the main component⁶⁰, being more consistent with the suggestion that low-mass AGB stars played this role. Such low-mass TP-AGB stars would operate at lower temperatures and with neutrons primarily coming from the (α, n) reaction on ^{13}C rather than ^{22}Ne . Modern calculations for these stars, although still parametric, reproduce the noble-gas observations on meteoritic silicon carbide in astonishing detail^{33,34}. Less success has been obtained from an alternative approach that tries to replicate the observed rather than the inferred pure s-process isotope patterns by exposing matter of Solar System composition to a small (~ 0.006 millibarn $^{-1}$) neutron dose³⁷.

The new type of calculation for AGB star s-process nucleosynthesis takes into account variations of neutron density and

temperature as prescribed by a stellar evolution calculation. This, in turn, predicts production of neutrons in a series of pulses each involving two bursts: a long one from the (α, n) reaction on ^{13}C , and a later, shorter one with the reaction on ^{22}Ne . One prediction is a finite contribution to isotopes such as ^{148}Nd and ^{134}Xe , which have traditionally been thought to have been produced exclusively by rapid neutron capture (r-process) without any s-process contribution. Again, this is supported by the silicon carbide data, which indicate contributions to these two isotopes at the level of a few per cent of the s-only isotopes^{56,61}. Analysis of interstellar silicon carbide thus provides information on details of the s-process that is impossible to gain by other means. Trying, for example, to decompose Solar System xenon or neodymium into the different nucleosynthetic components with comparable accuracy would be hopeless.

The AGB star theory: problems and numbers of stars. As I have said, those nucleosynthesis calculations for TP-AGB stars that reproduce many of the isotopic features in the SiC grains rely on just one stellar evolution model, which predicts the required neutron sources, whereas other evolution calculations do not (for a critical review see refs 36, 37). For the parameters needed to reproduce the trend of the s-process krypton composition with grain size, the trends they give for barium and strontium are opposite to those observed^{52,54}. It may be possible to remedy this situation because the krypton, strontium and barium results show different sensitivity to neutron dose and density, or by invoking different dominant sources for the noble gases and the other elements⁵⁹. But this leaves the problem of the apparent inability to reproduce the trend of the silicon isotopes, which, in a three-isotope plot of $\delta(^{29}\text{Si}/^{28}\text{Si})$ against $\delta(^{30}\text{Si}/^{28}\text{Si})$ define a line with a slope ~ 1.3 (Fig. 6) rather than the predicted ~ 0.5 (ref. 33).

A possible explanation might be that a variety of low-mass AGB stars of different metallicity contribute to the silicon carbide, the metallicities reflecting the effects of galactic chemical evolution^{40,62}. This might produce the observed slope (although the uncertainties are large), and s-processing proper within the stars would cause some scatter around the line (with a slope of ~ 0.5).

Another solution that has been offered⁶³ is the process of 'magnesium burning', which may occur in late stages of AGB stars and presumably could produce ^{29}Si and ^{30}Si in the right ratio. However, this would have to be in hot, intermediate-mass stars rather than low-mass stars, giving problems (as mentioned) in accounting for the s-process krypton results. Nevertheless, it has been suggested that a single, intermediate-mass AGB star, which possibly triggered the formation of the Solar System⁶⁴, may have contributed essentially all SiC near the forming sun⁶³.

Whether one star or a multitude was responsible for the 'typical' interstellar silicon carbide found in primitive meteorites cannot be determined from silicon isotope data alone, but if low-mass stars are responsible, a large number may be required. From simple statistical considerations, taking into account the amount of material expelled from an AGB star, the volume over which it is mixed and the volume of space that presumably would have been sampled by the protosolar cloud, it has been suggested that tens or up to 100 AGB stars may have contributed silicon carbide to the Solar System⁴⁰. The estimate has large uncertainties, however, and a value close to one cannot definitely be ruled out. A test might be to obtain a reliable estimate of the presolar age of the grains in question, a point that will be discussed in detail further below.

Interstellar diamond

Diamonds are two orders of magnitude more abundant in primitive meteorites than graphite and SiC (Table 1). They contain nitrogen with $\delta^{15}\text{N} = (-343 \pm 16) \text{‰}$ ($^{14}\text{N}/^{15}\text{N} = 414$)⁶⁵ and carry xenon-HL (Fig. 3), but the carbon isotope ratio itself ($^{12}\text{C}/^{13}\text{C} \approx 92$) is hardly unusual. This ratio (as well as the

nitrogen concentration) seems to vary with meteorite class, $\delta^{13}\text{C}$ ranging from -32 to -38‰ (ref. 65). The lack of more informative data may be due to the small average grain size of only $\sim 1.6 \text{ nm}$ diameter (mass-weighted; ref. 66) which means that a large number of grains are needed for any isotope analysis.

Xenon-HL was at first thought to have a fission origin, according to theories that concentrated on the better-characterized excesses of the heavy isotopes (Xe-H), and, as the spectrum did not match known fission spectra, a superheavy element was suggested as a progenitor (for example, ref. 67). But this did not explain the excesses of the light isotopes (L-Xe)⁶⁸, and with the discovery of isotopically unusual nitrogen⁶⁹ and the failure to find appropriate amounts of fission-produced Ba⁷⁰, the fission hypothesis was abandoned in favour of a nucleosynthetic origin.

For a long time practically the only study of possible nuclear sources for xenon-HL^{71,72} was one suggesting that Xe-H was produced during explosive nucleosynthesis in massive stars by neutron capture reactions (a 'mini r-process' intermediate between the classical s- and r-processes) and Xe-L was produced deeper in the same star (in the O-shell) by photodisintegration of a Ba/Xe seed, mixing with Xe-H later. But the data do not require an origin from the same star (despite the failure to separate xenon-H from xenon-L in the laboratory). Only about one in a million diamonds contains a xenon atom and there could be several populations of diamonds that we have been unable to separate from each other, of which one may carry Xe-H and another Xe-L.

Along similar lines, a recent study⁷³ investigates nucleosynthesis of heavy elements (silicon upwards) due to a neutron burst, as might occur in the outer helium-rich mantle of a massive star traversed by a supernova shock, with parameters adjusted to match the Xe-H pattern. To compare with these predictions, unfortunately, we only have recent data for barium and strontium⁷⁴ (apart from noble gas data), and these are not yet precise enough for a meaningful comparison.

A clue to how Xe-HL entered its carrier phase comes from the lack of evidence for enrichment of ^{129}Xe (which presumably would have entered as ^{129}I)⁷⁵ and barium^{53,74} relative to Xe-HL, as processes such as condensation or adsorption should favour chemically active barium and iodine over inactive xenon. Ion implantation might be a more promising process^{16,75}. This could have happened⁷⁶ if hot plasma from a supernova overtook a previously ejected carbon-rich dust shell, implanting trace elements, including Xe-HL, into the dust grains. It has been argued, however, that the diamonds and Xe-HL cannot have formed in the same star; instead, formation might have occurred in a binary star system⁷⁷. Specifically, this would mean synthesis of xenon-HL in a type I supernova and implantation during its eruption into diamonds formed in the outflow of the companion carbon star. Such a sequence of events would essentially decouple the isotopic structures seen in the light elements (C, N) from those in the heavier ones such as xenon. In another variation on this, the diamonds would have formed from pre-existing hydrocarbon grains around a supernova by annealing due to absorption of far-ultraviolet photons, with heavy ions and neutrals from the supernova outflow (including Xe-HL) implanted into the freshly synthesized grains²⁶.

Xenon-HL is not the only noble gas component carried by diamond. He and Ne-A2 (Table 1) of 'ordinary' isotopic composition are also present, and the isolation of very pure diamond has made it clear that xenon-HL is not even the only xenon component present. Rather, during heating, an additional component with roughly 'normal' isotopic composition is released at very low temperature⁷⁸. Electron energy-loss spectrometry has shown that the diamond residues actually consist of a mixture of diamond proper and some form of amorphous carbon, probably diamond-like hydrocarbon, a finding consistent with their apparent hydrogen content⁷⁹; the low-

temperature xenon component may be sited in hydrogen-rich surface layers. A third xenon component shows up during stepwise heating or stepwise combustion as slight deviations from mixing lines between xenon-HL and ordinary xenon^{80–83}. Its composition is not yet well defined, so the difficult task of determining a nucleosynthetic process for its origin has not been attempted.

Pre-solar phases and the early Solar System

Pre-solar graphite has so far only been directly identified in the Murchison meteorite, but on the basis of the occurrence of Ne-E(L) (Table 1), it would appear that it is present in other primitive meteorites. Pre-solar diamond and silicon carbide are more widespread, being present in carbonaceous as well as unequilibrated enstatite and ordinary chondrites (more primitive than type 3.7; refs 40, 84, 85). With increasing metamorphic type, the abundances of both SiC and diamond (and their noble gases) decrease, SiC more rapidly than diamond⁸⁵. If all meteoritic material started with the same endowment of pre-solar phases, this is surprising, as in the laboratory diamond burns and, in vacuum, releases its noble gases at lower temperature than SiC. Apart from the variations of the abundances of the interstellar grains with meteoritic type, some of their properties also seem to vary: the isotopic composition of diamond carbon, its nitrogen content⁶⁵, and the composition of the noble gases carried by SiC⁸⁶, the latter possibly related to differences in mean grain size.

Implications of survival. The fact that some interstellar grains have survived until today testifies that they never have been exposed to very high temperatures (especially not if coupled with high oxygen fugacity). Probably the most stringent temperature limits come from the persistence of Ne-E(L) in graphite and the isotopically ordinary ('normal') Xe component contained in diamonds because of the very low temperature at which these gases are released (significant release occurs already at <400 °C; for the latter component, though, an origin within the Solar System cannot be excluded.) Metamorphic temperatures experienced by the minerals in type 1 and 2 carbonaceous chondrites are certainly compatible with the constraints from the presence of these noble gas components and their carrier phases. But the temperatures needed to explain the differences in the abundance patterns of highly labile and moderately volatile elements among the bodies (meteorites as well as planetary) of the inner Solar System are clearly in excess of what graphite, diamond and the associated gases can have experienced.

In discussing this subject, one usually tacitly assumes that the interstellar grains were part of the Solar System during its formative stage. But, as they seem to occur as isolated grains within the meteorites⁸⁷, we do not know their whereabouts in the early Solar System and can say with certainty only that they were present in what are now the meteorites at the time of their compaction. That they were part of the Solar System before that will be hard to prove, but the case for it would be strengthened if they could be shown to be older than the Solar System. One way to show this would be to date them in a conventional way, but this seems near impossible; they do not carry significant quantities of any element usually used for age dating via radioactive decay (K, Rb, U, Sm), and as they came from exotic nucleosynthetic sources, we do not know with certainty the underlying isotopic compositions of the daughter elements. Another way would be to determine the duration of their pre-solar exposure to cosmic rays. If this age, added to their compaction age, exceeds the age of the Solar System, this would prove that they pre-dated the solar nebula (but not necessarily that they were present in it).

Pre-solar cosmic-ray exposure

An age for the cosmic-ray exposure of the pre-solar phases might also, in the case of SiC, put constraints on the possible

number of stellar sources. Of the three interstellar phases, only silicon carbide is susceptible to the analysis of its pre-solar cosmic-ray exposure age. This is because the rare isotope ²¹Ne is produced from silicon as a target element in about the same abundance as the other neon isotopes, so from its overabundance relative to the trapped neon composition, the duration of the cosmic-ray exposure can be calculated⁸⁸ (with correction for the ²¹Ne produced recently, during the lifetime of the meteoroid). Recent results⁶¹ indicate apparent pre-solar ages up to ~130 Myr, but there are complications.

First, with the amounts of cosmogenic ²¹Ne in the different grain size fractions determined from the deviations of data points from a mixing line between neon in the AGB He-shell and normal neon, the results for the different grain sizes depend strongly on the predictions for the composition of AGB star neon (although the relative results are probably less affected). Second, tiny grains lose a large fraction of cosmogenic nuclides produced within them by recoil, and experimental evidence on the extent of these losses would be helpful. Finally, only a few per cent of those SiC grains individually analysed contain detectable amounts of Ne-E(H)⁴⁷. It has been argued, therefore, that the rest may have lost, along with Ne-E(H), the neon from its pre-solar exposure, so the true exposure age may be much longer, in the range 1–2 Gyr. But it remains an open question whether Ne-E-free grains ever had any trapped neon, and therefore whether this conclusion is valid.

Future directions

The fact that surviving interstellar grains do exist should encourage further research, continuing the investigation of established phases as well as searching for other phases in meteorites.

Established phases. Some working models exist, but many open questions remain. All three established phases are carbonaceous and probably originate in carbon stars. The most complete isotopic information exists on SiC, for which isotopic features of the main and trace elements for most grains follow expected patterns for condensation products in the envelopes or winds of AGB stars. But problems do exist, and suggestions of origins in different carbon-rich stellar environments should not be lightly dismissed. Exotic grains of SiC seem to require the action of a supernova, possibly similar in some ways to what has been suggested for the diamonds (see above).

It may be fruitful to continue the search for other exotic grains of SiC, in addition to the already established grains 'X' and 'Y' (perhaps to be dubbed 'Z', but where would we go from there?). More heavy trace elements should be analysed for their signature of the s-process. Of special interest are those for which the compositions are sensitive to branchings and thus record the effects of temperature and neutron density during the process; examples are zirconium, molybdenum and gadolinium (recording the effects of branchings at ⁹⁵Zr and both ¹⁵¹Sm–¹⁵²Eu–¹⁵³Gd and ¹⁵⁴Gd; see ref. 89). Further progress may be expected from the use of more efficient and/or selective ionization techniques for mass spectrometry (for example, negative thermal ions or laser resonance ionization). The case of diamond has just been reopened by the realization of how complex the noble-gas inventory really is. Work on other trace elements, as done for SiC, may give insight into its source(s). Graphite, finally, has so far provided much less information than SiC or diamond. The best plan would be to search for it in other meteorite classes, and to investigate more of the minor and trace elements within it.

Other phases from carbon stars. Most would probably occur together with graphite, SiC and diamond, possibly in solid solution. Experimental evidence on the condensation of phases in the envelope of carbon stars suggests that SiC⁹⁰ (and/or graphite⁹¹) is the first of the major components to condense, and at lower temperatures other compounds including PAHs could condense on this⁹⁰. The chances of finding them either separately or within other interstellar grains (possibly in solid solution)

are probably highest for compounds that would condense at temperatures similar to or higher than the established interstellar grains. Examples are TaC, WC, NbC, ZrC, HfC and possibly Mo₂C, MoC, VC and Cr₃C₂ (ref. 91). TiC, which also belongs to this group, has actually been found within graphite (ref. 92) as well as within SiC (ref. 93). Most of these phases would probably be too small to be analysed individually for their isotopic composition.

Some nitrides (such as AlN, Si₃B₄) and sulphides (CaS) are also refractory under reducing conditions, but these would generally be more difficult to recover from the meteorites. Among these, Si₃N₄ has been found in the enstatite chondrite Quingzhen, but its isotopic composition (in Si, N, C) does not support a pre-solar origin⁹⁴. The situation may be different for ordinary chondrites⁹⁵, but evidence so far is only circumstantial.

Oxides. The other class of refractory compounds — stable under oxidizing rather than reducing conditions — is the refractory oxides, which have long been known to be the first phases to condense from a gas of solar composition and have been found in refractory Ca–Al-rich inclusions (CAIs) in primitive meteorites. As pointed out above, isotopic anomalies found in FUN and other CAIs probably represent reservoirs isotopically not equilibrated with the rest of the Solar System. This probably holds for the isotope abundance anomalies observed in a number of elements (H, N, O, Ti, Cr; (refs 4, 11, 96–99) throughout a meteorite, or, on a larger scale, varying from meteorite to meteorite or among meteorite classes. But this does not exclude the possible presence of small refractory interstellar oxide (or other) grains in the inclusions or elsewhere in the meteorites.

One corundum (Al₂O₃) grain has been found in the Orgueil meteorite, with an extraordinary inferred ²⁶Al/²⁷Al ratio of $\sim 9 \times 10^{-4}$, almost 20 times higher than the canonical value of 5×10^{-5} , possibly indicating a pre-solar origin¹⁰⁰; another has been found in the Murchison meteorite, with similar inferred ²⁶Al/²⁷Al and also highly anomalous oxygen ($\delta(^{17}\text{O}/^{16}\text{O}) = +1.072$; $\delta(^{18}\text{O}/^{16}\text{O}) = -244$ ‰; ref. 101). The high-temperature minerals corundum and hibonite (CaAl₁₂O₁₉) share with the carbonaceous grains a resistance to acids, making them easier to isolate. But some other possibly interesting phases do not, and isolating them will require a new approach not relying primarily on their chemical properties.

Anomalous nature of pre-solar grains. Investigation of the established pre-solar grains is of interest in its own right, but it is debatable whether general conclusions can be drawn about all interstellar grains from those that survived. For one, the surviving interstellar phases we know are, in all likelihood, not the 'typical' material that together with interstellar gas brought in the elements and their isotopes in the mix that defines the solar abundance pattern. This is obvious from their isotopic compositions which (with the exception of diamond carbon) are, even on average, different from the solar ones. Perhaps most important, diamonds carry abundant xenon-HL, which apparently has been created by a process ('mini' r-process) not at all necessary for the description of the solar abundances and which, therefore, may not have made a significant contribution to the Solar System at all.

And we know that the interstellar grains we now find in the meteorites now must be atypical simply because they did survive, in contrast to most grains and despite being carbonaceous in an oxidizing Solar System. Nevertheless, continued research on these (and, we hope, on new types) promises to yield exciting results for some time to come, providing useful clues to nucleosynthesis as well as to mixing, destruction and other processes in the interstellar medium and probably also the early Solar System. □

Ulrich Ott is at the Max-Planck-Institut für Chemie, Abteilung Isotopenkosmologie, Joh.-J. Becher-Weg 27, Postfach 55020, D-55128 Mainz, Germany.

- Burbidge, E. M., Burbidge, G. R., Fowler, W. A. & Hoyle, F. *Rev. mod. Phys.* **29**, 547–650 (1957).
- Clayton, D. D. *Principles of Stellar Evolution and Nucleosynthesis* (Univ. Chicago Press, 1983).
- Fowler, W. A. *Science* **226**, 922–935 (1984).
- Thiemens, M. H. in *Meteorites and the Early Solar System* (eds Kerridge, J. F. & Matthews, M. S.) 899–923 (Univ. of Arizona Press, Tucson, 1988).
- Mauersperger, K. *Geophys. Res. Lett.* **14**, 80–83 (1987).
- Hunten, D. M., Pepin, R. O. & Walker, J. C. G. *Icarus* **69**, 532–549 (1987).
- Suess, H. E. A. *Rev. Ast. Astrophys.* **3**, 217–234 (1965).
- Reynolds, J. H. & Turner, G. J. *Geophys. Res.* **69**, 3263–3281 (1964).
- Black, D. C. & Pepin, R. O. *Earth planet. Sci. Lett.* **6**, 395–405 (1969).
- Clayton, R. N., Grossmann, L. & Mayeda, T. K. *Science* **182**, 485–488 (1973).
- Lee, T. in *Meteorites and the Early Solar System* (eds Kerridge, J. F. & Matthews, M. S.) 1063–1089 (Univ. of Arizona Press, Tucson, 1988).
- Smith, V. V. in *Cosmic Abundances of Matter* (ed Waddington, C. J.) 200–223 (Am. Inst. Phys., New York, 1989).
- Iben, I. Jr in *Evolution of Stars: The Photospheric Abundance Connection* (eds Michaud, G. & Tutukov, A.) 257–274 (Kluwer, Dordrecht, 1991).
- Lewis, R. S., Anders, E. & Draine, B. T. *Nature* **339**, 117–121 (1989).
- Tang, M., Anders, E., Hoppe, P. & Zinner, E. *Nature* **339**, 351–354 (1989).
- Ott, U. in *Protostars and Planets III* (eds Levy, E. H., Lunine, J. & Matthews, M. S.) 883–902 (Univ. of Arizona Press, Tucson, 1993).
- Anders, E. & Zinner, E. *Meteoritics* **28** (in the press).
- Lewis, R. S., Srinivasan, B. & Anders, E. *Science* **190**, 1251–1262 (1975).
- Lewis, R. S., Tang, M., Wacker, J. F., Anders, E. & Steel, E. *Nature* **326**, 160–162 (1987).
- Bernatowicz, T. et al. *Nature* **330**, 728–730 (1987).
- Zinner, E., Tang, M. & Anders, E. *Nature* **330**, 730–732 (1987).
- Amari, S., Anders, E., Virag, A. & Zinner, E. *Nature* **345**, 238–240 (1990).
- Tielens, A. G. G. M. in *Carbon in the Galaxy: Studies from Earth and Space*, 59–111 (NASA Conf. Publ. 3061, 1990).
- Allamandola, L. J., Sandford, S. A., Tielens, A. G. G. M. & Herbst, T. M. *Astrophys. J.* **399**, 134–145 (1992).
- Allamandola, L. J., Sandford, S. A., Tielens, A. G. G. M. & Herbst, T. M. *Science* **260**, 64–66 (1993).
- Nuth, J. A., III & Allen, J. E. Jr *Astrophys. Space Sci.* **196**, 117–123 (1992).
- Amari, S., Lewis, R. S. & Anders, E. *Geochim. cosmochim. Acta* **57** (in the press).
- Mathis, J. S., Rimpl, W. & Nordsieck, K. H. *Astrophys. J.* **217**, 425–433 (1977).
- Virag, A. et al. *Geochim. cosmochim. Acta* **56**, 1715–1733 (1992).
- Hoppe, P., Amari, S., Zinner, E., Ireland, T. & Lewis, R. S. *Astrophys. J.* (submitted).
- Smith, V. V. & Lambert, D. L. *Astrophys. J. Suppl. Ser.* **72**, 387–416 (1990).
- Busso, M., Gallino, R., Lambert, D. L., Raiteri, C. M. & Smith, V. V. *Astrophys. J.* **399**, 218–230 (1992).
- Gallino, R., Busso, M., Picchio, G. & Raiteri, C. M. *Nature* **348**, 298–302 (1990).
- Gallino, R., Raiteri, C. M. & Busso, M. *Astrophys. J.* (in the press).
- Sackmann, I.-J. & Boothroyd, A. I. in *Evolution of Stars: The Photospheric Abundance Connection* (eds Michaud, G. & Tutukov, A.) 275–285 (Kluwer, Dordrecht, 1991).
- Arnould, M. in *Evolution of Stars: The Photospheric Abundance Connection* (eds Michaud, G. & Tutukov, A.) 287–297 (Kluwer, Dordrecht, 1991).
- Arnould, M. & Howard, W. M. *Meteoritics* **28** (in the press).
- Swart, P. K., Grady, M. M., Pillinger, C. T., Lewis, R. S. & Anders, E. *Science* **220**, 406–410 (1983).
- Amari, S., Zinner, E. & Lewis, R. S. *Lunar planet. Sci.* **XXII**, 19–20 (1991).
- Alexander, C. M. O. D. *Geochim. cosmochim. Acta* **57** (in the press).
- Tang, M. et al. *Geochim. cosmochim. Acta* **52**, 1221–1234 (1988).
- Fowler, W. A., Caughlan, G. R. & Zimmermann, B. A. A. *Rev. Astr. Astrophys.* **5**, 525–570 (1967).
- Caughlan, G. R. & Fowler, W. A. *Astrophys. J.* **136**, 453–464 (1962).
- Zinner, E., Amari, S., Anders, E. & Lewis, R. S. *Nature* **349**, 51–54 (1991).
- Ireland, T. R., Zinner, E. K. & Amari, S. *Astrophys. J.* **376**, L53–L56 (1991).
- Hoppe, P., Amari, S., Zinner, E. & Lewis, R. S. *Meteoritics* **27**, 235 (1992).
- Nichols, R. H. Jr, Hohenberg, C. M., Hoppe, P., Amari, S. & Lewis, R. S. *Lunar planet. Sci.* **XXIII**, 989–990 (1992).
- Amari, S., Hoppe, P., Zinner, E. & Lewis, R. S. *Astrophys. J.* **394**, L43–L46 (1992).
- Amari, S., Hoppe, P., Zinner, E. & Lewis, R. S. *Lunar planet. Sci.* **XXIII**, 27–28 (1992).
- Lewis, R. S. & Amari, S. *Lunar planet. Sci.* **XXIII**, 775–776 (1992).
- Ott, U. & Begemann, F. *Lunar planet. Sci.* **XXI**, 920–921 (1990).
- Prombo, C. A., Podosek, F. A., Amari, S. & Lewis, R. S. *Lunar planet. Sci.* **XXIII**, 1111–1112 (1992).
- Ott, U. & Begemann, F. *Astrophys. J.* **353**, L57–L60 (1990).
- Prombo, C. A., Podosek, F. A., Amari, S. & Lewis, R. S. *Astrophys. J.* (in the press).
- Zinner, E., Amari, S. & Lewis, R. S. *Astrophys. J.* **382**, L47–L50 (1991).
- Richter, S., Ott, U. & Begemann, F. *Lunar planet. Sci.* **XXIII**, 1147–1148 (1992).
- Käppeler, F., Beer, H. & Wisshak, K. *Rep. Prog. Phys.* **52**, 945–1013 (1989).
- Ott, U., Begemann, F., Yang, J. & Epstein, S. *Nature* **332**, 700–702 (1988).
- Lewis, R. S., Amari, S. & Anders, E. *Nature* **348**, 293–298 (1990).
- Busso, M., Picchio, G., Gallino, R. & Chieffi, A. *Astrophys. J.* **326**, 196–207 (1988).
- Lewis, R. S., Amari, S. & Anders, E. *Geochim. cosmochim. Acta* **57** (in the press).
- Gallino, R., Raiteri, C. M., Busso, M. & Matteucci, F. *Astrophys. J.* (submitted).
- Brown, L. E. & Clayton, D. D. *Astrophys. J.* **392**, L79–L82 (1992).
- Cameron, A. G. W. in *Protostars and Planets III* (eds Levy, E. H., Lunine, J. & Matthews, M. S.) 47–73 (Univ. of Arizona Press, Tucson, 1993).
- Russell, S. S., Arden, J. W. & Pillinger, C. T. *Science* **254**, 1188–1191 (1991).
- Fraundorf, P., Fraundorf, G., Bernatowicz, T., Lewis, R. & Tang, M. *Ultramicroscopy* **27**, 401–412 (1989).
- Anders, E., Higuchi, H., Gros, J., Takahashi, H. & Morgan, J. W. *Science* **190**, 1262–1271 (1975).
- Manuel, O. K., Hennecke, E. W. & Sabu, D. D. *Nature Phys. Sci.* **240**, 99–101 (1972).
- Lewis, R. S., Anders, E., Wright, I. P., Norris, S. J. & Pillinger, C. T. *Nature* **305**, 767–771 (1983).
- Lewis, R. S., Anders, E., Shimamura, T. & Lugmair, G. *Science* **222**, 1013–1015 (1983).
- Heymann, D. & Dziczkaniec, M. *Proc. 10th Lunar planet. Sci. Conf.* 1943–1959 (1979).
- Heymann, D. & Dziczkaniec, M. *Meteoritics* **15**, 15–24 (1980).
- Howard, W. M., Meyer, B. S. & Clayton, D. D. *Meteoritics* **27**, 404–412 (1992).
- Lewis, R. S., Huss, G. R. & Lugmair, G. *Lunar planet. Sci.* **XXII**, 807–808 (1991).
- Lewis, R. S. & Anders, E. *Astrophys. J.* **247**, 1122–1124 (1981).
- Clayton, D. D. *Proc. lunar planet. Sci.* **B12**, 1781–1802 (1981).
- Jørgensen, U. G. *Nature* **332**, 702–705 (1988).
- Tang, M. & Anders, E. *Geochim. cosmochim. Acta* **52**, 1245–1254 (1988).
- Bernatowicz, T. J., Gibbons, P. C. & Lewis, R. S. *Astrophys. J.* **359**, 246–255 (1990).

80. Lewis, R. S. & Anders, E. *Lunar planet. Sci.* **XX**, 679–680 (1988).
81. Nichols, R. H. Jr. *et al. Geochim. cosmochim. Acta* **55**, 2921–2936 (1991).
82. Verchovsky, A. B., Ott, U., Russell, S. S., Fisenko, A. V. & Shukolyukov, Yu. A. *Lunar planet. Sci.* **XXIII**, 1467–1468 (1992).
83. Huss, G. R. & Lewis, R. S. *Meteoritics* **28** (in the press).
84. Alexander, C. M. O'D., Arden, J. W., Ash, R. D. & Pillinger, C. T. *Earth planet. Sci. Lett.* **99**, 220–229 (1990).
85. Huss, G. R. *Nature* **347**, 159–162 (1990).
86. Ott, U., Lohr, H. P. & Begemann, F. *Meteoritics* **26**, 382 (1991).
87. Alexander, C. M. O'D., Swan, P. & Walker, R. M. *Nature* **348**, 715–717 (1990).
88. Tang, M. & Anders, E. *Astrophys. J.* **335**, L31–L34 (1988).
89. Beer, H. *Astrophys. J.* **375**, 823–832 (1991).
90. Frenklach, M., Carner, C. S. & Feigelson, E. D. *Nature* **339**, 196–198 (1989).
91. Lodders, K. & Fegley, B. Jr *Meteoritics* **27**, 250–251 (1992).
92. Bernatowicz, T. J., Amari, S., Zinner, E. K. & Lewis, R. S. *Astrophys. J.* **373**, L73–L76 (1991).
93. Bernatowicz, T. J., Amari, S. & Lewis, R. S. *Lunar planet. Sci.* **XXIII**, 91–92 (1992).
94. Alexander, C. M. O'D., Prombo, C. A., Swan, P. D. & Walker, R. M. *Lunar planet. Sci.* **XXII**, 5–6 (1991).
95. Lee, M. R., Russell, S. S., Arden, J. W. & Pillinger, C. T. *Meteoritics* **27**, 248–249 (1992).
96. Zinner, E. in *Meteorites and the Early Solar System* (eds Kerridge, J. F. & Matthews, M. S.) 956–983 (Univ. of Arizona Press, Tucson, 1988).
97. Prombo, C. A. & Clayton, R. N. *Science* **230**, 935–937 (1985).
98. Sugiura, N. & Hashizume, K. *Earth planet. Sci. Lett.* **111**, 441–454 (1992).
99. Rotaru, M., Birck, J. L. & Allègre, C. J. *Nature* **358**, 465–470 (1992).
100. Huss, G. R., Hutcheon, I. D., Wasserburg, G. J. & Stone, J. *Lunar planet. Sci.* **XXIII**, 563–564 (1992).
101. Nittler, L. R., Walker, R. M. & Zinner, E. *Lunar planet. Sci.* **XXIV**, 1087–1088 (1993).
102. McCulloch, M. T., & Wasserburg, G. J. *Astrophys. J.* **220**, L15–L19 (1978).

ACKNOWLEDGEMENTS. I thank V. V. Smith for providing Fig. 2, S. Amari and E. K. Zinner for providing Figs 4–6, M. Arnould, F. Begemann, T. J. Bernatowicz, R. Gallino, P. Hoppe, A. Jorissen and E. K. Zinner for discussions and R. S. Lewis and I. P. Wright for constructive reviews and the MacDonnell Center for the Space Sciences for its hospitality during revision of the manuscript.

Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1

Jerry H. Brown^{*}, Theodore S. Jardetzky, Joan C. Gorga[†],
Lawrence J. Stern^{*}, Robert G. Urban, Jack L. Strominger & Don C. Wiley^{*†}

Department of Biochemistry and Molecular Biology, ^{*} Howard Hughes Medical Institute, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA

[†] Department of Pediatrics, Children's Hospital of Pittsburgh, 6130 Rangos Research Building, 3705 Fifth Avenue, Pittsburgh, Pennsylvania 15213, USA

The three-dimensional structure of the class II histocompatibility glycoprotein HLA-DR1 from human B-cell membranes has been determined by X-ray crystallography and is similar to that of class I HLA. Peptides are bound in an extended conformation that projects from both ends of an 'open-ended' antigen-binding groove. A prominent non-polar pocket into which an 'anchoring' peptide side chain fits is near one end of the binding groove. A dimer of the class II $\alpha\beta$ heterodimers is seen in the crystal forms of HLA-DR1, suggesting class II HLA dimerization as a mechanism for initiating the cytoplasmic signalling events in T-cell activation.

CLASS II human leukocyte antigens (HLA) are cell surface $\alpha\beta$ heterodimers ($M_r \sim 60,000$) expressed primarily by specialized antigen-presenting cells such as macrophages, dendritic cells and B lymphocytes. Class II HLA presents peptides derived from extracellular antigens (refs 1, 2 but see refs 3, 4) to helper T cells as part of the mechanism for identifying foreign antigens and producing an immune response.

Class II HLA is polymorphic in the human population⁵ and different types selectively bind different sets of peptides^{1–3}. The HLA polymorphism appears to be responsible for variations in the immune response of different individuals to different antigens, and may contribute to the susceptibility to disease and autoimmune disorders^{6,7}.

Crystals have been grown from a number of class II HLA-DR (ref. 8) allelic variants after purification by immunoaffinity chromatography in detergent complexes and removal of the transmembrane region and cytoplasmic domain by papain digestion⁹ and from a soluble construct of HLA-DR1 expressed in insect cells¹⁰.

The three-dimensional structure of the extracellular portion of HLA-DR1 (DRA/DRB1*0101) has now been determined using three different crystal forms. In each crystal form, DR1 crystallizes as a dimer of the DR1 $\alpha\beta$ heterodimer. The class II structure is similar to that of the class I HLA molecule¹¹ and to a hypothetical model of the class II binding site based on the class I structure¹². Structural differences from class I HLA include an altered position of the immunoglobulin-like β_2

domain of HLA-DR1 relative to that of the α_3 domain of class I HLA, and detailed changes in the peptide binding site so that its ends are more open. The peptide-binding site observed in the crystal structure contains electron density from the collection of peptides copurified with DR1 from human lymphoblastoid cells⁴. These peptides are bound as straight, extended chains that project out of both ends of the site. At one position of the peptide a side chain extends into a distinct pocket in the binding site. The main chain of the bound peptide contacts conserved, hydrogen-bonding residues along the HLA helices.

The presence of a parallel dimer of the $\alpha\beta$ heterodimers in crystals suggests the capacity of class II HLA molecules to dimerize before, or more probably during, recognition by helper T cells. Such a dimer of dimers could lead to an increased affinity for the CD4 coreceptor and to the crosslinking of T cell receptors to initiate cytoplasmic signalling pathways.

Structure determination

The three-dimensional structure of the extracellular portion of DR1 was determined to 3.3 Å resolution by combining X-ray diffraction data from three different types of crystals containing DR1: (1) Papain cleaved DR1 from human B lymphoblastoid cells carrying a mixture of endogenous peptides^{8,9}; (2) the same DR1-peptide complex plus the 28K superantigen *Staphylococcus aureus* enterotoxin B (T.S.J. *et al.*, unpublished results); and (3) a complex of soluble DR1 expressed in insect cells with an antigenic peptide from the influenza virus haemagglutinin, HA 306–318 (ref. 10) (Table 1 legend).

An initial model of the polypeptide chain was determined from

[†]To whom correspondence should be addressed.

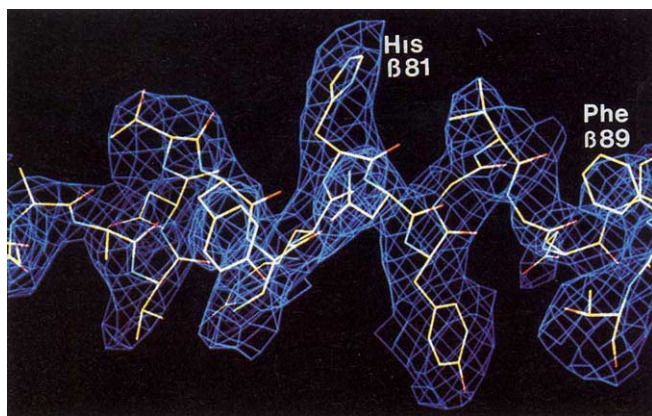


FIG. 1 Sixfold iteratively averaged 3.3 Å 'omit' electron density of the β_1 domain helix. About 83% of the polypeptide side chains show electron density of similar quality. Residues α 38, 46, 47, 50, 57, 126 and β 59, 64, 189 show no side-chain electron density, whereas residues α 1, 51, 52, 53, 132, 143, 181, β 42, 66, 68, 69, 70, 71, 72, 94, 108, 144 presently have ambiguous electron density.

an electron density map at 4.2 Å resolution, using phases from a single isomorphous heavy-atom derivative (0.12 mM K_2HgI_4 , 0.3 mM KI) and iterative, fourfold, real-space averaging¹³ of the first two crystal forms. Fourfold averaging was possible because each of the two crystals contained two crystallographically independent copies of the α - β heterodimer related by a non-crystallographic 2-fold symmetry axis (the symmetry axis in the dimer of heterodimers, Fig. 5a). Phases were improved by model building and iterative sixfold real-space averaging about the non-crystallographic 2-fold symmetry axes of all three crystal forms.

Real-space averaged 'omit'¹⁴ maps, now at 3.3 Å resolution, show excellent density for 83% of the polypeptide side chains (Fig. 1). Nine side chains show no electron density. Seventeen residues have ambiguous electron density, including β -chain residues 68–72 which have been provisionally modelled as helical. Though electron density is observed to emanate from each of the three oligosaccharide attachment sites (most extensively at α 118), it has not yet been interpreted as a molecular model. Refinement of DR1 in the three crystal forms is in progress.

Structural comparison with class I HLA

Overall the structure of DR1 is similar to that of class I HLA (Fig. 2a). The two α -chain domains, α_1 and α_2 , of DR1 superimpose closely on the corresponding α_1 domain and the β_2 -microglobulin subunit, respectively, of class I HLA. The two β -chain domains, β_1 and β_2 , of DR1 superimpose on the α_2 and less closely on the α_3 domains, respectively, of class I HLA. The β_2 domain of DR1 is offset relative to the HLA-A2, class I α_3 domain by about 15° (Fig. 2a), but the packing between the class II immunoglobulin domains still resembles the packing in class I more than the packing in an antibody.

The peptide-binding groove of both class I and class II HLA is composed of eight strands of antiparallel β -sheet as a floor and two antiparallel helical regions as the sides. The similarity between the two structures extends even to the location of β -bulges in the β -strands and the locations of kinks in the helices (Fig. 2b). With the exception of the sequence alignment in strand 4 (residues α 37 to α 43) of the α_1 domain, a hypothetical model of class II (ref. 12) based on the class I HLA-A2 structure is correct. (A number of segments deemed unpredictable based on sequences and the class I structure have a different structure in class II antigens.)

Differences in the helical regions of class II HLA relative to class I may determine why class I HLA binds primarily short, nonameric peptides¹⁵, whereas class II HLA binds longer, 15–

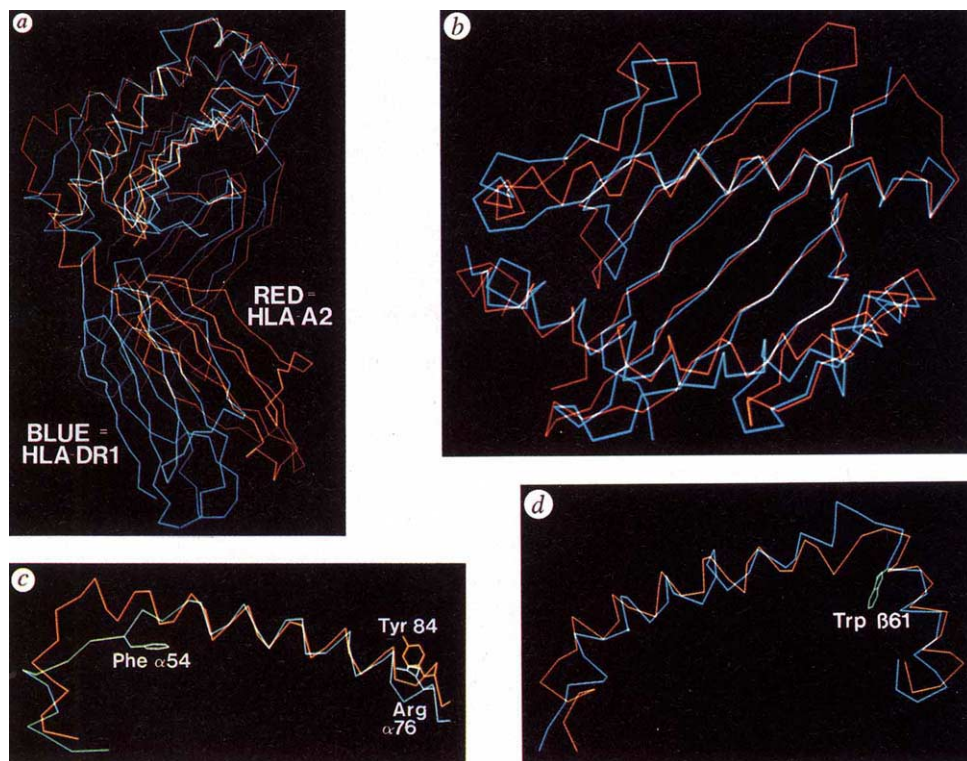


FIG. 2 Class II HLA-DR1 (blue) superimposed on class I HLA-A2 (red). a, One DR1 $\alpha\beta$ heterodimer and one HLA-A2 molecule³⁴. The β_2 DR1 domain is offset from its homologue the α_3 HLA-A2 domain by about 15°. b, 'Top' view shows the overall similarity of the peptide binding sites. HLA-A2 α_1 and DR1 α_1 domains' helices are at the top, those of HLA-A2 α_2 and DR1 β_1 are at the bottom. c, 'Side' view of α_1 domain helical regions showing differences at both ends. On the left (also seen in b) DR1 has an extended chain where the HLA-A2 has helical residues. On the right, DR1 is about 2.0 Å lower (at Arg α 76) than HLA-A2 Tyr 84. d, 'Side' view of the helical regions of the DR1 β_1 and HLA-A2 α_2 domains. DR1 has a more prominent kink and Trp β 61 superimposes almost exactly on HLA-A2 Trp 147. α -Carbon coordinates were superimposed by the program FITATOM (W. Kabsch) based on the central six β strands and two long helices of the peptide-binding site.

In addition to these differences in secondary structure between the class I and class II binding groove, a number of specific amino-acid side chains that 'close' the ends of the class I site are not present in the class II site. The conserved tyrosines (7, 59, 171) that bind the peptide N terminus and the conserved residues (Y84, K146, T143) that bind the peptide C terminus in the class I structure^{17, 19} are not found in class II sequences. Furthermore, a nearly conserved tryptophan (167) and salt bridge (Glu 55 to Arg 170) which 'close' the 'left' (Fig. 2) end of the class I groove²⁰ are absent in class II sequences. The significantly smaller side chain (glycine or valine) at β 86 compared to the class I homologue tyrosine 171 helps create a deep pocket in the DR1 groove surrounded by predominately non-polar residues of both the α_1 and β_1 domains (left in Fig. 3). This pocket, which provides the binding site for a peptide 'anchor' side chain in DR1, is composed of class II residues homologous to class I residues beyond the 'left' end of the class I binding groove.

The distribution of conserved and polymorphic positions in the class II peptide-binding site suggests how conserved class II resi-

Several clusters of polymorphic residues, spaced across the binding site, appear to provide the mechanism by which different types of class II HLA selectively bind peptides of different sequences²³⁻²⁶. The first cluster (α 31, 52; β 86, 89, 90) appears to provide variability to a peptide side chain pocket composed

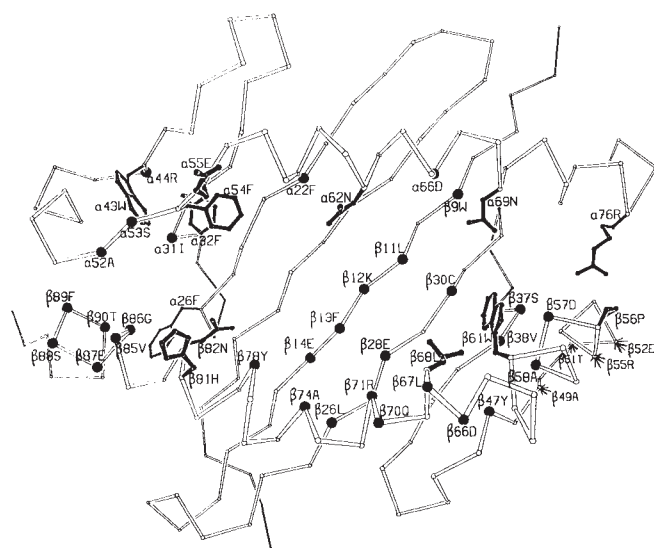


FIG. 3 The class II α_1 and β_1 domains, showing all polymorphic positions (dark circles, POL > 2¹²) and those conserved positions facing the antigen-binding groove (all atoms of side chain shown). Amino-acid residues, in one letter code, and positions are given for DR1 (DRA/DRB1*0101). The antigen-binding groove of DR1 includes the side chains of residues 7, 9, 11, 22, 24, 26, 31, 32, 43, 51, 53, 54, 55, 58, 62, 65, 68, 69, 72, and 76 of the α chain, and residues 9, 11, 13, 28, 30, 32, 37, 38, 47, 56, 60, 61, 65, 68, 70, 71, 74, 78, 81, 82, 85, 86, 88, and 89 of the β chain. Residues whose side chains are in the dimer interface (see text and Fig. 5) are indicated by 'starbursts'.

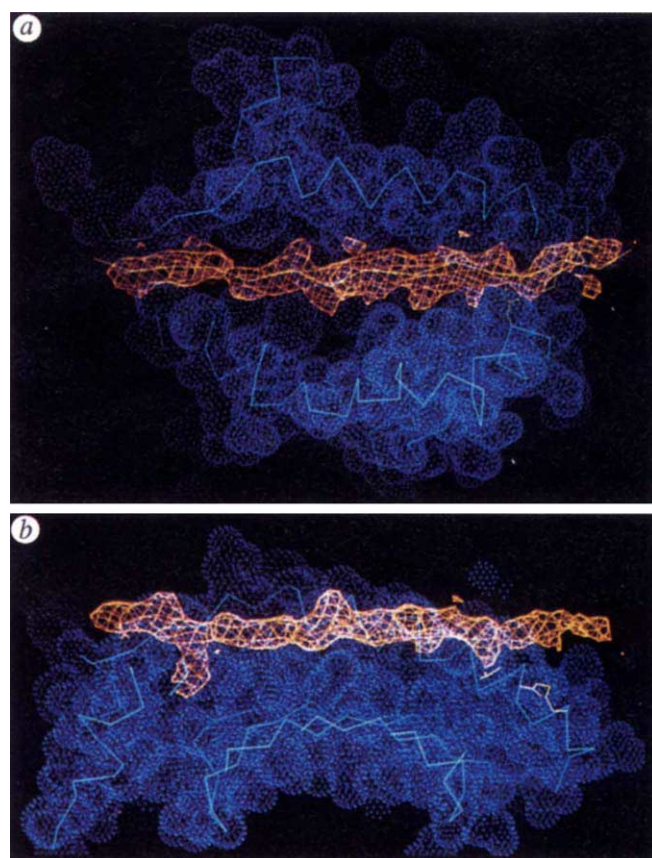


FIG. 4. Peptides bound to crystalline DR1 from human cells. Red, Electron density of the collection of endogenous bound peptides; Blue, van der Waals surface of DR1. a, 'Top' view, probably as seen by T cells; α_1 domain helix on top, β_1 domain helix on bottom; peptide N terminus at left. An extended, 15 amino-acid long, α -carbon, peptide backbone is placed in the electron density for scale. The side chains shown at the right-hand end of the site are arginine $\alpha 76$ and aspartic acid $\beta 57$ which form a salt bridge under the exiting peptide. b, 'Side' view, β_1 domain helix removed for clarity, β -sheet shown below and α_1 domain helix shown behind peptide electron density. The prominent downward facing peptide side chain electron density near the left end of the peptide is located about one peptide residue to the left of where the N terminus of peptides bound to class I HLA are positioned. Electron density map is fourfold iteratively averaged, 3.3 Å resolution, using DR1-LG2 and DR1-SEB crystal forms (see Table 1 legend) and $2F_o - F_c$ coefficients. Density within 1.5 Å of provisional peptide atoms is included.

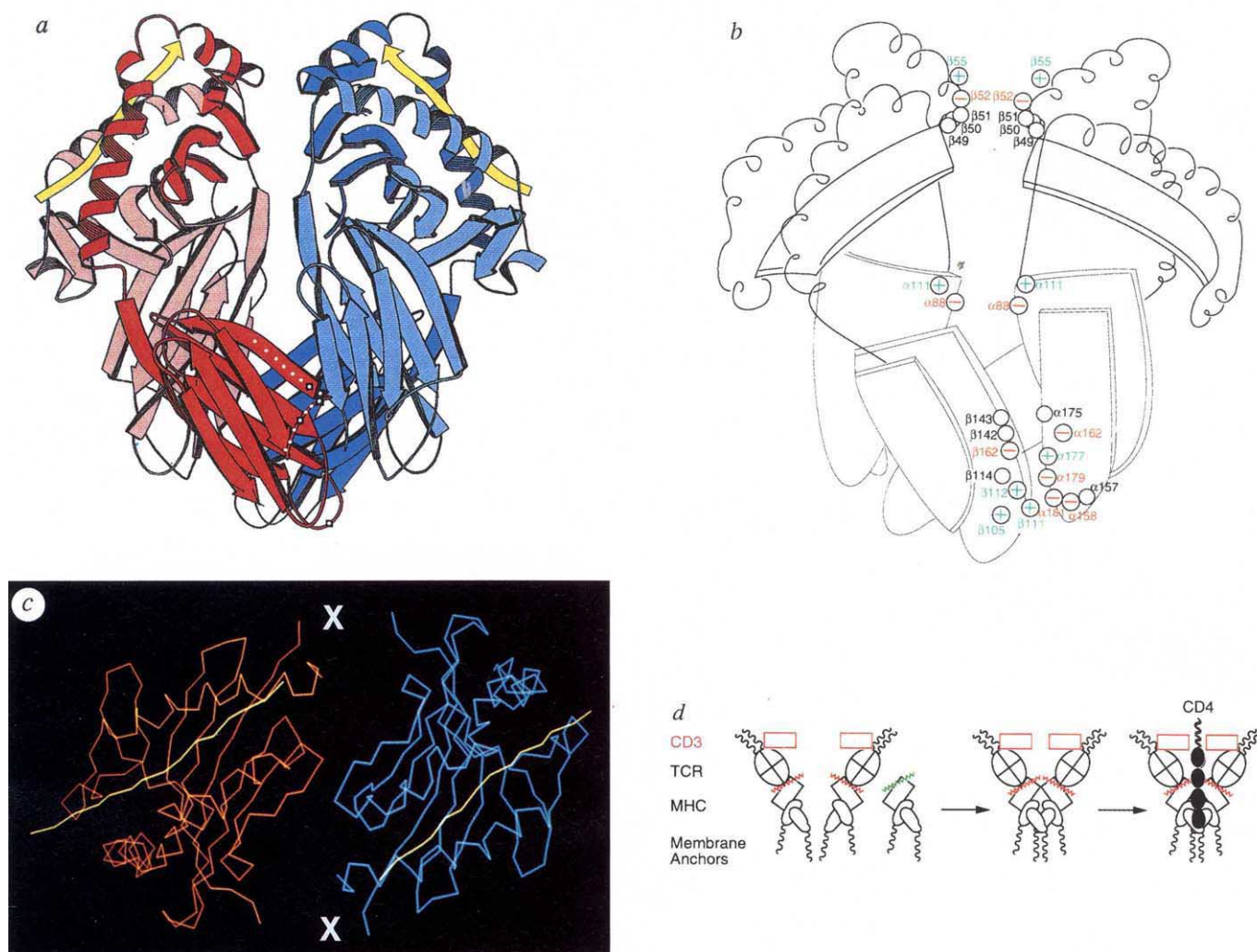


FIG. 5 DR1 crystallizes as a dimer of the $\alpha\beta$ heterodimer. *a*, One DR1 molecule is shaded in red, the other in blue, peptides are yellow. The α -chains are the lighter shades, the β -chains are darker. The width of the dimer shown is 76 Å; the height is 73 Å. β_2 residues implicated in CD4 binding by mutagenesis⁴⁰ (□), by peptide binding⁴¹ (○). α_2 domain residues near the β_2 domain residues that have been implicated in CD4 binding are α 116, 118, 120, 122, 125, 127, 129, 164, 166, 168, 171, 172 and 173. *b*, Residues which make contact (<3.5 Å) in current, unrefined model across the DR1 dimer interfaces are indicated. Red, Asp or Glu; Blue, Arg, Lys, His. The second dimer-related interface at the bottom is not shown. A total of 1,293 Å² of one DR1 heterodimers solvent accessible surface is buried by the other, larger than the interface between antibody and protein antigen⁶⁸. Surface areas in DR1

were obtained using ACCESS with a probe radius of 1.4 Å (written by M. D. Handschumacher and F. M. Richards), and may slightly change upon refinement. *c*, The dimer of antigen-binding domains as seen from the top. The closest approach of the bound portions of the two peptides (in yellow) is ~30 Å, the furthest distance being about 80 Å. 'X' marks the probable locations of the CD4 coreceptors. *d*, Hypothetical assembly pathway for an antigen-specific, cell-cell interaction complex of class II HLA, T-cell receptor, and the coreceptor CD4 (see text). Further aggregation may occur if CD4 dimerizes⁶⁹ (J. Wang and S. C. Harrison, personal communication). (An alternative model where CD4 binds before TCR has similar properties but is less consistent with CD4-independent T-cell recognition.) The box marked CD3 is intended to symbolize the CD3, ζ , and associated components of the T-cell receptor.

of mostly conserved (α 26, 32, 43, 54) non-polar residues in class II HLA (Fig. 3). Three other polymorphic clusters are found at three, five and about seven or eight residues further along the bound peptide, composed of β 13, 26, 28, 71, 74, 78; α 66, β 9, 11, 30 and β 37, 38, 58. The clustering of the murine I-E and I-A equivalents of β 71 and β 74 with β 28 has previously been suggested from mutual compensatory effects of mutations at those residues on peptide binding and T cell response²⁷. The structure is also consistent with earlier observations of the effect of allelic specific residues on α - β chain pairing^{28,29}.

The polymorphic residue aspartic acid β 57 in DR1 appears to form a salt bridge with a conserved arginine α 76 (Fig. 3) under the exiting peptide (see below). Aspartic acid β 57 in the HLA-DQ isotype has been found to correlate with protection against insulin-dependent diabetes mellitus (IDDM) in humans,

as HLA-DQ allelic products with side chains of β 57 incapable of forming a salt bridge are linked to susceptibility to IDDM^{7,30}. In the non-obese diabetic mouse alterations at β 57 in I-A^k have been implicated in insulinitis³¹⁻³³, but other genes also appear to be involved in generating disease³¹. In addition to a possible effect on peptide binding, polymorphism at β 57 may affect the stability or structure of the dimer of $\alpha\beta$ heterodimers because nearby residues β 52 and β 55 form part of the interface of the dimer (see below).

Structure of bound endogenous peptides

The electron density in the binding groove (Fig. 4), which represents the collection of endogenous peptides from human lymphoblastoid cells, is straight and thin as expected for a peptide in an extended conformation. This density indicates that about

TABLE 1 Data collection statistics

X-ray Temp.	Protein (source)	Space group (Z*)	a, b, c (Å)	Deriv.	Resolution R_{merge} comp(red)†	Resolution R_{merge} comp(red)†	ΔF	r.m.s. f_h/E ‡
GX-13 4 °C	DR1(LG-2)	C222 ₁ (2)	96.7 112.6 209.2	native	100–4 Å 0.117 74% (1.5)	4.5–4.2 Å 0.222 60% (1.4)		
GX-13 4 °C	DR1(LG-2)	C222 ₁ (2)	96.6 112.6 209.2	K ₂ HgI ₄	15–4.2 Å 0.099 73% (1.2)	4.4–4.2 Å 0.198 63% (1.1)	0.20 (15–4.2 Å)§	1.49
GX-13 –165 °C	DR1(LG-2) SEB	P2 ₁ 2 ₁ 2 ₁ (2)	95.7 115.9 150.4	native	30–3.4 Å 0.101 70% (1.44)	3.6–3.4 Å 0.24 51% (1.47)		
GX-13 –165 °C	DR1(LG-2) SEB	P2 ₁ 2 ₁ 2 ₁ (2)	95.9 117.1 150.8	K ₂ HgI ₄	30–3.2 Å 0.077 83% (4.0)	3.4–3.2 Å 0.229 76% (2.23)	0.185 (30–3.5 Å)§	1.3
CHESS –170 °C	DR1(LG-2)	C222 ₁ (2)	96.7 114.4 216.8	native	15–2.7 Å 0.082 90% (2.4)	3.4–3.2 Å 0.168 95% (2.7)		
CHESS –170 °C	DR1(LG-2) SEB	P2 ₁ 2 ₁ 2 ₁ (2)	95.0 114.7 149.8	native	20–2.8 Å 0.057 88% (3.3)	3.0–2.8 Å 0.224 89% (3.5)		
CHESS –170 °C	DR1(SF9) HA306–318	P4 ₃ 2 ₁ 2 (2)	95.4 95.4 245.7	native	20–2.7 Å 0.068 93% (3.2)	3.0–2.9 Å 0.221 94% (3.2)		
Averaging statistics								
R^s			4.2 Angstrom DR1(LG-2) 0.23	4-fold DR1/SEB 0.23		3.3 Angstrom DR1(LG-2) 0.31	6-fold DR1/SEB 0.30	DR1(SF9) 0.30
Mean-phase								
Difference								
Overall			75.3°	68.7°		56.8°	67.5°	51.1°
Last cycle			2.8°	3.3°		3.2°	4.8°	3.1°

DR1 and a number of other subtypes and isotypes of human class II histocompatibility antigens have been crystallized⁸ after purification by immunoaffinity chromatography in detergent complex and removal of the transmembrane region and cytoplasmic domain by papain digestion⁹. Crystals of DR1 isolated from the human B lymphoblastoid cell line LG-2 (DR1–LG2) suitable for high-resolution study were initially grown at pH 4.0 and later at pH 5.5–6.0 from 27% PEG 8000, 0.1 M glycine, 0.01% Na₂S₂O₈, 0.05 M MES. Crystals of the superantigen SEB in 1:1 complex with DR1–LG2 were grown using the same human-cell derived DR1 mixed with SEB (Sigma) further purified by gel filtration and will be described further elsewhere (T.S.J. et al., manuscript in preparation). DR1 expressed as a soluble heterodimer in SF9 insect cells (DR1–SF9), purified as an empty molecule, and subsequently complexed with the influenza virus 13-mer peptide HA 306–318 crystallized from 15% PEG 8000, 0.02% Na₂S₂O₈, 100 mM sodium acetate, pH 4.6¹⁰. Crystals of DR1–LG2 and the DR1–SEB complex were soaked overnight in 0.12 mM K₂HgI₄, 0.3 mM KI. Native and derivative X-ray data (first four sets of rows) were collected with a Xentronics area detector using an Elliot GX-13 X-ray source focused with Franks mirrors⁶², then integrated and scaled using the BUDDHA⁶³ and CCP4 packages. Two heavy-atom positions per asymmetric unit were found to be 27.5 Å apart by analysis of difference Patterson maps of both crystals. Electron density maps were calculated at 6 Å resolution using single isomorphous replacement (SIR) phases for DR1–LG2 and SIR with anomalous scattering for DR1–SEB. A relationship between the electron density in the two crystal forms was found by a one-dimensional, real-space correlation search along the Hg to Hg vector in the two crystals using the program CYLINDER (written by P. Metcalf). The maps from the two crystals were then averaged. An envelope could be constructed around the DR1 portion of the electron density from the average map, and the remaining electron density in the DR1–SEB crystals was assigned to an SEB envelope. Thirteen cycles of iterative, real-space averaging within the DR1 envelopes and solvent flattening outside of DR1 and DR1–SEB in the two crystal forms were done using CCP4 and Joy of skewing⁶⁴. Input for the first cycle of averaging include observed structure factors (F_{obs}) and single isomorphous replacement phases (ϕ_{SIR}) for DR1–LG2 and F_{obs} and ϕ_{SIR} /anomalous for DR1–SEB. Input for cycles 2–4 include F_{obs} and combined phases (ϕ_{comb})¹³. Input for cycles 5–8 included combined structure factors (F_{comb})⁶⁴ and ϕ_{comb} . Input for the final 5 cycles included Sim-weighted $2F_{\text{obs}} - F_{\text{calculated}}$ coefficients and calculated phases. This procedure yielded an electron density map from which the domain structure was evident for two, $\alpha\beta$ heterodimers of DR1 per asymmetric unit in each crystal. Using the 2-fold noncrystallographic rotational symmetry within each crystal's asymmetric unit, the above averaging procedure was repeated, but now real-space averaged 4-fold within the DR1 envelopes, 2-fold within the SEB envelopes, and solvent flattened outside the envelopes at 4.2 Å resolution (left side of averaging statistics). The DR1 polypeptide pathway and many large side chains were evident at this point. The Hg derivative was seen to label the only free cysteine (β 30). Higher resolution, more complete X-ray data were then collected on phosphor image plates at CHESS (and integrated using DENZO (Zbyszek Otwinowski, personal communication)) from crystals that had been soaked in cryoprotectant (row sets 5–7) and mounted in nichrome (DR1–SEB) or glass (DR1–LG2 and DR1–SF9) loops⁶⁵ and flash frozen to –165 °C in a stream of cold nitrogen. A partially refined model of DR1 (against synchrotron-cryo DR1–LG2) was fit and refined as rigid domains using X-PLOR version 2.1 (written by A. Brünger) into the 3 or 3.5 Å resolution cryo-X-ray data of the other two crystals. For each of the three crystal forms, eight 'omit models' were constructed¹⁴, each missing about one-eighth of the DR1 model. The DR1–LG2 'omit' models were subjected to a simulated annealing refinement procedure to reduce bias⁶⁶. 'Omit maps' were calculated from each 'omit' model. Eight cycles of sixfold, iterative, real-space averaging of the electron density within the DR1 envelopes of both monomers in the 'omit' maps of all three crystals (right side of averaging statistics) yielded the current 3.3 Å (Fig. 1) map. A model has been built using FRODO⁶⁷.

* Z, Number of DR1 heterodimers per asymmetric unit.

† Comp. % completeness = $100 \times (\text{number of independent reflections}) / (\text{maximum number theoretical})$, red, redundancy = $(\text{number of measurements}) / (\text{number of independent reflections})$. CHESS = Cornell High Energy Synchrotron Source, GX-13 = Elliot rotating anode.

‡ f_h , Heavy-atom structure factor; E the residual lack of closure.

§ Resolution range used for heavy-atom derivative statistics, ΔF and r.m.s. f_h/E .

$$R_{\text{Merge}} = \frac{\sum_{hkl} \sum_{\text{obs}} |f_{\text{obs}}^{hkl} - \langle f_{\text{obs}}^{hkl} \rangle|}{\sum_{hkl} \sum_{\text{obs}} f_{\text{obs}}^{hkl}}, \quad \Delta F = \frac{\sum_{hkl} |F_{\text{native}} - F_{\text{derivative}}|}{\sum_{hkl} F_{\text{native}}}, \quad R = \frac{\sum_{hkl} |F_{\text{observed}} - F_{\text{calculated from average map}}|}{\sum_{hkl} F_{\text{observed}}}$$

|| SEB source: SEB used in these frozen crystals was a gift from M. Sax, VA Medical Center, Pittsburgh, PA. The same crystals were also obtained from SEB purchased from Toxin Technology Inc., Sigma, or purified from an SEB producing strain of *S. aureus*.

15 peptide residues contact HLA-DR, consistent with the average length of the most abundant peptides bound to DR1 from LG2 cells⁴. The electron density is not weak in the middle as observed in class I HLA-A2³⁴ and HLA-Aw68³⁵ (but not HLA-B27¹⁷) where peptides of different length appear to be accommodated by bulging out of the site near the middle^{17,19,35}. Instead it extends out of both ends of the site, consistent with the observation that the endogenous peptides of class II HLA vary in length from 12–24 amino acids and have 'ragged' ends^{1,4}.

The current DR1 model is based on averaging the 3.3 Å resolution data of different crystal types and is not crystallographically refined, so no detailed model for the bound peptides has been constructed. (A more complete analysis of peptide contacts is in preparation at higher resolution for the complex between DR1 and influenza virus haemagglutinin peptide, HA 306–318 (L.J.S. *et al.*, unpublished data). But the electron density from LG2 cell-derived DR1 (Fig. 4b, left end) indicates the position of one prominent peptide side chain that appears bound in a hydrophobic pocket near residues α 26, 31 and β 86. This seems to be the position of the dominant 'anchoring' peptide residue predicted at position 308 of the influenza HA (306–318) peptide both by binding studies of substituted peptide sequences and by the high-affinity binding of a simplified peptide where all but two residues were made alanine³⁶. The size of this peptide residue side chain seems to parallel the size of the non-polar pocket as modulated by the Gly–Val dimorphism at β 86^{37,39}. These observations and the homologous role observed for tryptophan β 61 in class II and tryptophan 147 in class I (see above) suggest that peptides are bound with their N termini to the left (in Fig. 3), an orientation homologous to that in class I sites. This has been confirmed for one peptide HA306–318 complexed with HLA-DR1, which is currently being refined to 2.7 Å resolution (L.J.S. *et al.*, unpublished data). The clarity of the peptide electron density may in part be due to the predominance of a few peptides in DR1 from LG2 cells⁴.

A dimer of $\alpha\beta$ heterodimers

Because a dimer (Fig. 5a) is observed in DR1 crystals, it may represent an intrinsic property of DR1 (see next section). There are four interfaces between DR1 molecules in the dimer (Fig. 5b). Two small interfaces are near the 2-fold rotational symmetry axis that relates the molecules (top two in Fig. 5b), β_1 49–55 and α_2 88, 111. In the most membrane-distal interface, five β_1 domain residues of one molecule contact the same residues in the β_1 domain of the second molecule, burying 305 Å² of solvent accessible surface on each monomer (see also Fig. 3). Four of these residues are conserved in all DR alleles. Two salt bridges from glutamic acid β 52 to arginine β 55 on the other molecule are conserved in DR, but are replaced by proline and leucine residues in some DP and DQ variants. The second interface is simply two salt bridges between glutamic acid 88 and lysine 111 of the second domain of the conserved DR1 α chain. This interface is small, burying 75 Å² of surface per monomer.

The bottom two interfaces are identical and are located about 12 Å from the dimer rotation axis at the membrane proximal end of the dimer (Fig. 5b). Each is composed of 7 residues on the β_2 domain of one molecule in contact with 7 residues on the α_2 domain of the other molecule and buries surfaces of about 900 Å². Most of the residues are conserved in class II HLA sequences including those participating in three salt bridges between α_2 domain residues glutamic acid 158, histidine 177, glutamic acid 179 and β_2 domain residues histidine 111, glutamic acid 162 and histidine 112.

CD4 binding site

The T-cell coreceptor molecule, CD4, is thought to bind to the β_2 domain of DR1 based on mutagenesis that indicated CD4 affinity changes when β_2 domain residues 110, 137, 140, 141 and 142 were altered⁴⁰ and based on the observation that peptides with the sequence of β_2 domain residues 134–148 (and less so,

138–152) could bind to CD4⁴¹ (white symbols in Fig. 5a). If the DR1 dimer is present during T-cell recognition, the X-ray structure suggests that a region on the α_2 domain of the second DR1 molecule in the dimer (light blue domain near white symbols, Fig. 5a) might also contact CD4, potentially increasing the affinity of the DR1–CD4 interaction over that possible between CD4 and a single DR1 molecule. (α_2 domain positions near the β_2 CD4 site are listed in Fig. 5a legend.) Some of the β_2 domain positions identified by mutagenesis as important for CD4 binding are in (β 141, 142) or near (β 110 for which the affinity effect is only partial⁴⁰) the dimer interface raising the possibility that some of the observed CD4 affinity changes may have been due to destabilization of a class II HLA dimer. Viewed from the 'top' (Fig. 5c), the DR1 dimer presents two peptide binding site surfaces presumably for recognition by T-cell receptors and two slots between them (X in Fig. 5c). The presumptive CD4-binding sites described above are located behind the Xs in Fig. 5c.

Conclusion

HLA class II versus class I peptide binding. The class II DR1 structure shows that the peptide-binding site is open at both ends (Fig. 4). Long peptides, at least 15 residues, are seen to be bound in an extended conformation, projecting out of both ends of the site. This contrasts with peptides bound to class I HLA molecules where mostly 9-mers bind with extended, but kinked, conformations, and the N and C termini of the peptides are sequestered in conserved, peptide-terminal binding sites^{17–19,35}. In the mechanism by which MHC molecules bind many peptides tightly, the role of clusters of conserved residues in class I HLA binding of peptide termini^{17,35,42} appears to be replaced in class II HLA by conserved residues (β 82N, α 62N, α 69N, β 61W, α 76R) spaced so as to bind at intervals (Fig. 3) to the extended main chain of bound peptide, providing a peptide side chain independent component to the affinity.

In class I MHC molecules polymorphic residues in the binding site form pockets which accommodate peptide side chains and which form the basis for allele-specific peptide binding^{17,19,34,35,43}. In class II DR1 one prominent peptide side chain at about the third position of a 15-mer identifies a critical binding pocket, and other pockets are inferred from the locations of polymorphic residues and the shape of the electron density representing multiple peptides complexed with DR1 (Figs 3 and 4). The critical binding pocket of DR1 lies near one end (left in Figs. 3 and 4) of the class II binding groove, before the location where the class I groove begins. The location of this critical pocket, the orientation of the peptide, and the extended conformation of the peptide are consistent with some of the previous studies of peptide interactions with class II proteins^{36,44,45}.

Possible implications of a class II HLA dimer of dimers. Two observations suggest that the dimer of DR1 $\alpha\beta$ heterodimers discovered in the three crystals studied is suitable for a physiological role. (1) In the dimer two DR1 molecules are parallel to each other so that all four C termini face one direction (down in Fig. 5a) and both peptide-binding sites face the opposite direction (up in Fig. 5a) as might be expected for glycoproteins on a membrane surface. (2) The two peptide-binding regions are oriented so that two molecules the size of the $\alpha\beta$ T-cell receptor could contact these binding sites simultaneously.

The possibility that the dimer is a crystallization artefact is not ruled out. It is also possible that a dimer stable only at high protein concentrations in solution has been stabilized by crystallization. Whether stabilization conditions might exist physiologically to take advantage of this capacity to form dimers is discussed below. Whether the dimer exists *in vivo* can be tested by site-directed mutagenesis. If the observed dimer of DR1 molecules is physiologically important, its structure suggests that mutations in the α_2 domain may affect CD4 binding (see Fig. 5a legend) and that mutations in the dimer interfaces (Fig. 5b) may affect both CD4 binding and T-cell activation.

The capacity of DR1 molecules to form a parallel dimer suggests the possibility that the observed dimerization might provide part of the mechanism for initiating intracellular signalling pathways both in T cells by crosslinking two T-cell receptors^{46,47} and, in antigen presenting cells, to induce the expression of costimulating molecules^{48–51}. Dimerization of cell-surface receptors with the concomitant dimerization of their cytoplasmic domains is an established mechanism for signal initiation in some hormone receptor interactions, for example that of human growth hormone⁵². Indeed, T-cell receptors can be activated by crosslinking with a divalent antibody but are not activated by the cognate monovalent Fab fragment⁵³. But hydrodynamic and other measurements provide no evidence that DR1 dimerizes in solution^{8,10,54} (although detergent-soluble DR1 migrates anomalously on gel filtration⁵⁵), even at the acidic conditions (pH 4.0–6.0) used in crystallization (L.J.S., unpublished data).

Low affinity in the dimer interaction may be an important feature for T-cell activation. Unless a mechanism exists for loading both peptide-binding sites of a single dimer with the same peptide, dimers preexisting on antigen-presenting cells would be expected to contain two different peptides, only one being specific for a given T cell. Weakly interacting DR1 molecules would exchange more readily, allowing T-cell receptor recognition a role in the assembly of an activation complex.

If the T-cell receptor binds to monomeric DR1 molecules and affects dimerization, many mechanisms for initiating signalling can be generated. For one example (Fig. 5d), assume that the $\alpha\beta$ T-cell receptor (and its associated CD3 and ζ subunits) can also form weak dimers like those of DR1, then before DR1–TCR binding neither DR1 nor TCR dimers would be stable

(Fig. 5d, left). Once DR1 molecules with a specific peptide are bound by a TCR, each such complex would be able to form a stable dimer by using both the weak dimerization sites of DR1 and of the TCR molecules simultaneously, achieving stability by a chelating effect (Fig. 5d, middle)⁵⁶. The DR1 dimer structure further suggests that the affinity for the coreceptor CD4 could be greater after dimerization due to contacts with both monomers, suggesting both that adhesion contributed by CD4 would be increased after TCR-dependent dimer assembly, and that signalling through CD4⁵⁷ could be triggered by its association with this complex (Fig. 5d, right). Concentrations sufficient to drive the assembly process (Fig. 5d) may be found in the synapse-like contacts between T cells and their target cell.

Some peptides that bind to class II MHC but engage the T-cell receptor weakly or differently do not initiate full T-cell signalling and therefore act as pure antagonists or mixed agonist/antagonists^{58–60}. Such MHC–TCR complexes might be expected to assemble into the interaction complex proposed above (Fig. 5d, right), but could destabilize it below the threshold required to stimulate complete intracellular signalling.

A synergistic, multisubunit assembly process, where each interaction increases the affinity of subsequent interactions, would have the potential to amplify an initially small, but specific, interaction between the $\alpha\beta$ chains of TCR and HLA into a stable, antigen-specific complex that bridges both sides of the T-cell membrane and the antigen-presenting cell membrane (Fig. 5d, right). This mechanism substitutes direct, stable molecular contacts to initiate full intracellular signalling for previous models that invoked statistical increases in the local density of TCR–HLA complexes to lead to T-cell activation^{47,61}. □

- Rudensky, A. Y., Preston-Hurlburt, P., Hong, S.-C., Barlow, A. & Janeway, C. A. Jr *Nature* **353**, 622–626 (1991).
- Hunt, D. F. et al. *Science* **256**, 1817–1820 (1992).
- Chicz, R. M. et al. *J. exp. Med.* **178**, 27–47 (1993).
- Chicz, R. M. et al. *Nature* **358**, 764–768 (1992).
- Kappes, D. & Strominger, J. L. A. *Rev. Biochem.* **57**, 991–1028 (1988).
- Tiwari, J. L. & Terasaki, P. I. *HLA and Disease Associations* (Springer, New York, 1985).
- Todd, J. A. et al. *Science* **240**, 1003–1009 (1988).
- Gorga, J. C., Brown, J. H., Jardetzky, T., Wiley, D. C. & Strominger, J. L. *Res. Immun.* **142**, 401–407 (1991).
- Gorga, J. C., Horejsi, V., Johnson, D. R., Raghupathy, R. & Strominger, J. L. *J. biol. Chem.* **262**, 16087–16094 (1987).
- Stern, L. J. & Wiley, D. C. *Cell* **68**, 465–477 (1992).
- Bjorkman, P. J. et al. *Nature* **329**, 506–512 (1987).
- Brown, J. H. et al. *Nature* **332**, 845–850 (1988).
- Bricogne, G. *Acta crystallogr.* **A32**, 832–847 (1976).
- Sussman, J. L., Holbrook, S. R., Warrant, R. W., Church, G. M. & Kim, S.-H. *J. molec. Biol.* **123**, 607–630 (1978).
- Falk, K., Rötzschke, O., Stevanovic, S., Jung, G. & Rammensee, H.-G. *Nature* **351**, 290–296 (1991).
- Gorga, J. C. *Crit. Rev. Immun.* **11**, 305–335 (1992).
- Madden, D. R., Gorga, J. C., Strominger, J. L. & Wiley, D. C. *Cell* **70**, 1035–1048 (1992).
- Silver, M. L., Guo, H.-C., Strominger, J. L. & Wiley, D. C. *Nature* **360**, 367–369 (1992).
- Fremont, D. H., Matsumura, M., Stura, E. A., Peterson, P. A. & Wilson, I. A. *Science* **257**, 919–927 (1992).
- Bjorkman, P. J. et al. *Nature* **329**, 512–518 (1987).
- Griffith, I. J. et al. *J. exp. Med.* **167**, 541–555 (1988).
- Peccoud, J., Dellabona, P., Allen, P., Benoist, C. & Mathis, D. *EMBO J.* **9**, 4215–4223 (1990).
- Sette, A. et al. *J. Immun.* **143**, 3289–3294 (1989).
- Sidney, J. et al. *J. Immun.* **149**, 2634–2640 (1992).
- Hammer, J., Takacs, B. & Sinigaglia, F. *J. exp. Med.* **176**, 1007–1013 (1992).
- Geluk, A. et al. *J. Immun.* **149**, 2864–2871 (1992).
- Racioppi, L., Ronchese, F., Schwartz, R. H. & Germain, R. N. *J. Immun.* **147**, 3718–3727 (1991).
- Braunstein, N. S. & Germain, R. N. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2921–2925 (1987).
- Braunstein, N. S., Germain, R. N., Loney, K. & Berkowitz, N. J. *Immun.* **145**, 1635–1645 (1990).
- Horn, G. T., Bugawan, T. L., Long, C. M. & Erlich, H. A. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6012–6016 (1988).
- Miyazaki, T. et al. *Nature* **345**, 722–724 (1990).
- Slattery, R. M. et al. *Nature* **345**, 724–726 (1990).
- Lund, T. et al., *Nature* **345**, 727–729 (1990).
- Saper, M. A., Bjorkman, P. J. & Wiley, D. C. *J. molec. Biol.* **219**, 277–319 (1991).
- Guo, H.-C. et al. *Nature* **360**, 364–368 (1992).
- Jardetzky, T. S. et al. *EMBO J.* **9**, 1797–1803 (1990).
- Busch, R., Hill, C. M., Hayball, J. D., Lamb, J. R. & Rothbard, J. B. *J. Immun.* **147**, 1292–1298 (1991).
- DeMotz, S., Barbey, C., Corradin, G., Amoroso, A. & Lanzavecchia, A. *Eur. J. Immun.* **23**, 425–432 (1993).
- Newton-Nash, D. K. & Eckels, D. D. *J. Immun.* **150**, 1813–1821 (1993).
- König, R., Huang, L.-Y. & Germain, R. N. *Nature* **356**, 796–798 (1992).
- Cammara, G. et al. *Nature* **356**, 799–801 (1992).
- Madden, D. R., Gorga, J. C., Strominger, J. L. & Wiley, D. C. *Nature* **353**, 321–325 (1991).
- Garrett, T. P. J., Saper, M. A., Bjorkman, P. J., Strominger, J. L. & Wiley, D. C. *Nature* **342**, 692–696 (1989).
- Krieger, J. I. et al. *J. Immun.* **146**, 2331–2340 (1991).
- Kropshofer, H., Böhlinger, I., Max, H. & Kalbacher, H. *Biochemistry* **30**, 9177–9187 (1991).
- Klinman, N. R. *J. exp. Med.* **136**, 241–260 (1972).
- Janeway, C. A. Jr et al. *Cold Spring Harb. Symp. quant. Biol.* **54**, 657–666 (1989).
- Koulova, L., Clark, E. A., Shu, G. & DuPont, B. J. *exp. Med.* **173**, 759–762 (1991).
- Nabavi, N. et al. *Nature* **360**, 266–268 (1992).
- Schwartz, R. H. *Cell* **71**, 1065–1068 (1992).
- Watts, T. H., Alaverdi, N., Wade, W. F. & Linsley, P. S. *J. Immun.* **150**, 2192–2202 (1993).
- DeVos, A. M., Ullrich, M. & Kossiakoff, A. A. *Science* **255**, 306–312 (1992).
- Kaye, J. & Janeway, C. A. Jr *J. exp. Med.* **159**, 1397–1412 (1984).
- Roche, P. A., Marks, M. S. & Cresswell, P. *Nature* **354**, 392–394 (1991).
- Springer, T. A., Kaufman, J. F., Terhorst, C. & Strominger, J. L. *Nature* **268**, 213–218 (1977).
- Jencks, W. P. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4046–4050 (1981).
- Glaichenhaus, N., Shastri, N., Littman, D. R. & Turner, J. M. *Cell* **64**, 511–520 (1991).
- DeMagistris, M. T. et al. *Cell* **68**, 625–634 (1992).
- Alexander, J. et al. *J. Immun.* **150**, 1–7 (1993).
- Racioppi, L., Ronchese, F., Matis, L. A. & Germain, R. J. *exp. Med.* **177**, 1047–1060 (1993).
- Matsui, K., et al., *Science* **254**, 1788–1791 (1991).
- Harrison, S. C. *J. Appl. Crystallogr.* **1**, 84–89 (1968).
- Blum, M., Metcalf, P., Harrison, S. C. & Wiley, D. C. *J. appl. Crystallogr.* **20**, 235–242 (1987).
- Stuart, D. & Artymuk, P. *Acta crystallogr.* **A40**, 713–716 (1984).
- Teng, T.-Y. *J. Appl. Crystallogr.* **23**, 387–391 (1990).
- Hodel, A., Kim, S.-H. & Bronger, A. T. *Acta crystallogr.* **A48**, 851–858 (1992).
- Jones, T. A. *J. Appl. Crystallogr.* **11**, 268–272 (1978).
- Colman, P. M. *Adv. Immun.* **43**, 99–132 (1988).
- Kwong, P. D. et al. *Proc. natn. Acad. Sci. U.S.A.* **87**, 6423–6427 (1990).

ACKNOWLEDGEMENTS. We thank P. Klimovitsky, A. Haykov and M. Frayser for technical assistance; K. Svenson, assisted by M. Pietras, for the large-scale production of tissue culture cells; and other members of our group and the CHESS synchrotron staff for assistance with X-ray data collection. We acknowledge conversations with R. Germain who reached similar conclusions about how a dimer might not preexist, but only be stabilized by TCR binding. J.H.B. is supported by the Howard Hughes Medical Institute, T.S.J. by the Cancer Research Institute; J.C.G. by the Juvenile Diabetes Foundation International. L.J.S. by a Damon Runyon–Walter Winchell Cancer Fellowship and the Howard Hughes Medical Institute; R.G.U. by the Irvington Institute; J.L.S. by the NIH. D.C.W. is an investigator with the Howard Hughes Medical Institute. Preliminary, unrefined coordinates of DR1 are available from the authors.

Two sources of diffuse X-ray emission from the Galactic Centre

M. Markevitch, R. A. Sunyaev & M. Pavlinsky

Institute for Space Research (IKI), Profsoyuznaya 84/32,
Moscow 117810, Russia

THE weak extended X-ray source at the Galactic Centre has so far lacked a reasonable explanation. Measurements of this roughly elliptical source from the Ginga satellite revealed¹ emission in the 6.7-keV line of ionized iron, indicating that the X-rays originate in an optically thin plasma. But to account for the hard X-ray spectrum, this plasma needs to be very hot — too hot, in fact, to be confined by the gravitational potential of the Galactic Centre². We have recently shown³ that the morphology of the emitting region changes at energies above 11 keV: the source becomes extended in the galactic plane, resembling the distribution of the molecular gas clouds in this region. Here we report the detection of a pronounced absorption feature in the emission spectrum in the energy range 8–11 keV. This result, combined with the changing spatial distribution, suggests that the high-energy emissions arise from the scattering of X-rays from the nearby compact sources by the dense molecular clouds. As no comparable absorption feature is seen at lower energies, the softer X-ray emissions may still be understood in terms of thermal emission from a plasma, but the required temperature is no longer unreasonably high.

A source of weak extended emission in the Galactic Centre region was discovered in soft X-rays by the Imaging Proportional Counter on the Einstein satellite⁴ and studied in all subsequent X-ray experiments. The first harder energy image of this source (4–15 keV at 12' resolution) was obtained during the Spacelab 2 mission⁵. Measurement from the satellite Ginga revealed emission in the 6.7-keV line of ionized iron from that region¹, suggesting that the observed flux may be bremsstrahlung from an optically thin plasma heated to several kiloelectronvolts. Later, the TTM instrument on board Mir/Kvant obtained a spectrum of this source in the 2–25 keV energy band (assuming a fixed elliptical brightness distribution and simultaneously solving for intensities of the diffuse source and the bright galactic bulge sources within the 16° field of view⁶). This spectrum is consistent with bremsstrahlung with a temperature of 12.7 ± 0.4 keV. Such a high temperature is more than an order of magnitude too high for a plasma to be confined in the Galactic Centre gravitational well (see, for example, ref. 2).

In 1990, the ART-P telescope on board the satellite Granat obtained 5' resolution images of this region in several energy bands over the 3–30 keV interval^{8,3}. The ART-P telescope has a fully coded 3.6° field of view, which eliminates the ambiguity in the reconstructed image and therefore the need to assume the shape of the region of emission. This is the first time that direct mapping of the diffuse source has been performed in different energy intervals above 3 keV. The results reported here are derived by accumulating the seven observations with a total effective exposure (corrected for collimator vignetting) of 58,500 s in the position of Sgr A. Figure 1 presents the resulting smoothed maps of the source in several energy intervals of interest. The observed properties of the source are discussed in detail in ref. 3.

The angular resolution of the images is sufficient to allow separation of the diffuse component from the bright projected compact sources (labelled in Fig. 1, top left) that are responsible for about two-thirds of the 3–22 keV flux from the 1° diffuse emission region. Figure 2a shows a spectrum of the diffuse component from the 1° × 1.5° ellipse (centred 10' from Sgr A and tilted at 30° with respect to the galactic plane, see Fig. 1, top right, omitting the circles containing A1742–294, 1E1743.1–

2843 and Sgr A, and that of the sum of these sources. Although the data on the diffuse emission are sparse, its spectrum is apparently unusual. A prominent absorption feature above 8 keV (obviously interpreted as photoabsorption mainly by neutral iron) is present in the diffuse spectrum, whereas it is much weaker in the spectrum of the compact sources. This indicates that the absorption occurs in the source of the diffuse emission itself rather than on the way to the observer. This absorption feature is more significant in the spectrum of the central bright 30' diameter part of the source (Fig. 2b). The energy resolution of the telescope is not sufficient to allow the study of the iron line in the diffuse spectrum.

We must carefully quantify the reality of the absorption feature because the source is weak and the statistics are poor. To do this, we fitted the spectrum shown in Fig. 2b with a power law or bremsstrahlung spectrum photoabsorbed in a neutral gas with solar abundances. This model does not fit the spectrum ($\chi^2=19.5$, with eight experimental points and three fitted parameters giving five degrees of freedom), apparently because the dip in the spectrum at 8–13 keV requires much stronger absorption than that necessary to fit at lower energies. An assumption of a 'warm' absorbing gas transparent for photons of energy $E < 7-8$ keV (with all elements but iron ionized) does not help, because such a dense material would itself emit too much. It may be seen in Fig. 1, however, that the diffuse emission has a different structure, and probably a different cause, at low and high energies. Below 8 keV (Fig. 1, top right), the source is roughly elliptical, in agreement with previous experiments. At higher energies (Fig. 1, bottom right), the source is elongated in the direction of the galactic plane and shows many features in common with the distribution of giant clouds of molecular gas. For example, enhancements in the X-ray brightness are seen at the positions of Sgr A, Sgr B2 and possibly Sgr C cloud complexes (see ref. 9 for a review). It is therefore appealing to suggest that the harder-band emission originates in this gas. It has been noted earlier that the galactic molecular gas may be visible in X-rays due to the electron scattering of photons from bright sources illuminating the gas (for example refs 7, 14). A simple estimate^{3,5} shows that a diffuse flux comparable to that we observe may arise from such illumination of the molecular clouds by the nearby compact X-ray sources. The scattering of X-ray photons by the electrons in hydrogen atoms occurs without a loss of energy (coherently) at $E \lesssim 2$ keV, whereas at higher photon energies both coherent and incoherent (Compton) scattering with ionization occurs. Taken together, these channels yield a scattering cross-section equal to that of Thomson scattering on free electrons (see, for example, ref. 10), and we refer to this sum as 'Thomson scattering'.

The scattered emission must be weak at lower energies because of the strong photoabsorption (since the photoionization depth τ_{phot} of a gas with solar abundances is greater than or comparable with that of Thomson scattering for $E \lesssim 12-14$ keV). We attempted to fit the five harder energy bins of the spectrum in Fig. 2b separately from the softer bins, using the same simple model. The fit is good ($\chi^2 = 0.3$, with (5–3) degrees of freedom). The F test (which is used to determine whether the data require an additional parameter of the model to be introduced (see, for example, ref. 11) shows that $\tau_{\text{phot}} > 0$ must be included as an additional parameter at $> 99\%$ confidence. (If increasing energy spectra are allowed, then the confidence level is 98%.) This means that the observed absorption feature is statistically significant. The value of χ^2 as a function of two fitted parameters (steepness of the incident spectrum and absorption column) is shown in Fig. 3, from which a lower limit on τ_{phot} may be found¹¹. The cut-off at energies below 13 keV requires a very strong absorption, $N_{\text{H}} > 5 \times 10^{23} \text{ cm}^{-2}$ ($\tau_{\text{phot}} > 1.5$ at $E=7.1$ keV) at a confidence level of 99% for all non-increasing incident energy spectra, with the best-fit value even greater. A value of the photoabsorption depth typical for the compact sources in

this region is an order of magnitude lower¹². This may be understood if both the scattering and the absorption that we observed occur in compact dense gas clouds, which are unlikely to obscure the background sources. If such clouds give the main contribution to the scattered emission, then its spectrum should have a photoabsorption typical of the clouds rather than that averaged over the corresponding area in the sky — that is, we see the clouds with high absorption depth and do not see other

regions because they do not scatter.

Although the applied model is too simplified (and the poor statistics make it unreasonable to test a more adequate model with scattering and absorption occurring in the same gas), it is evident that the softer portion of the emission scattered in the dense clouds would not escape such an opaque gas. Therefore, the observed elliptical source at energies below 8 keV (the three lower energy bins in Fig. 2) should have another cause. For

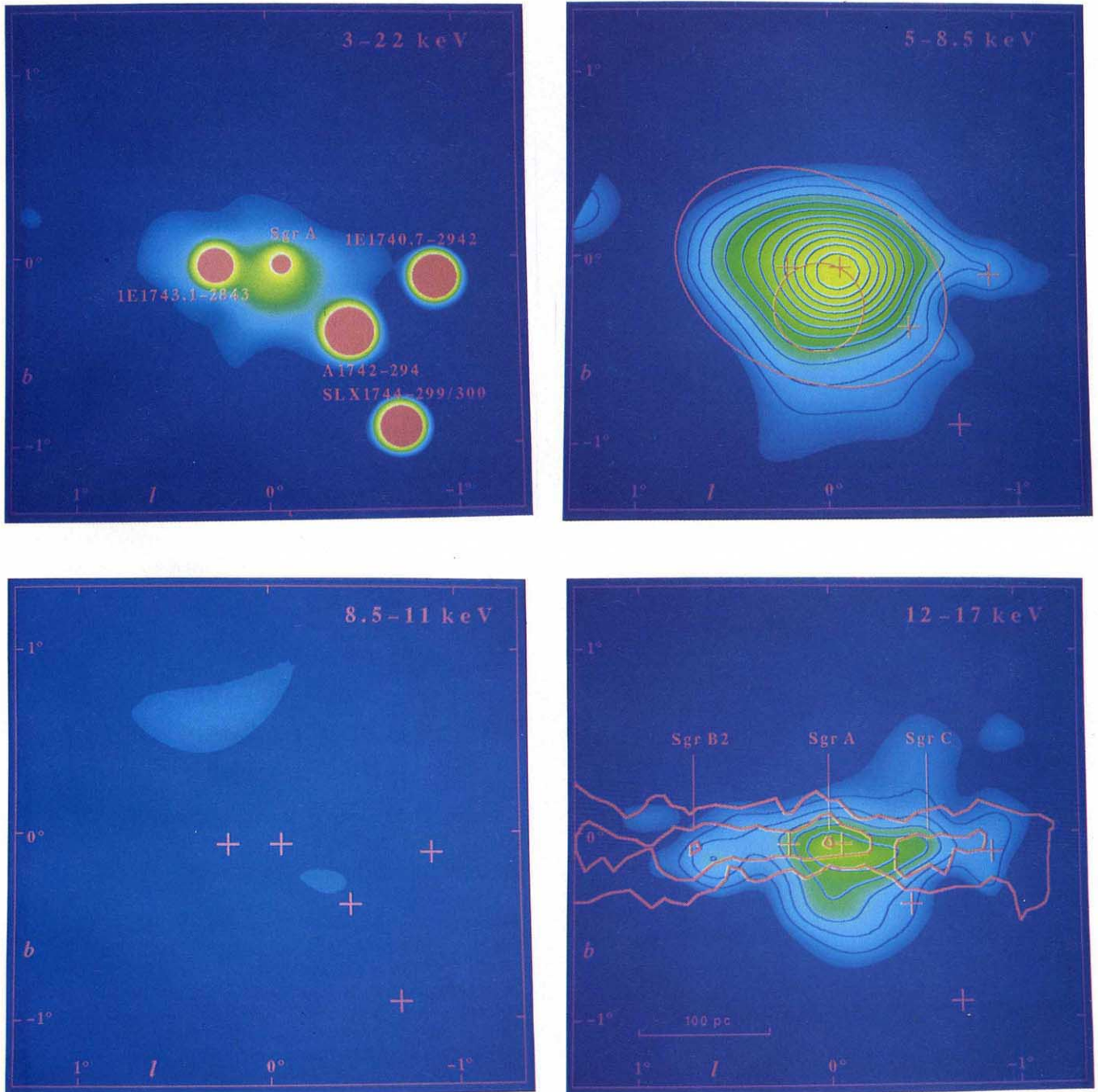


FIG. 1 Top left: The Galactic Centre image in the 3–22 keV energy band. The image with point sources removed was smoothed to improve the significance of the diffuse emission. The point sources were then put back with gaussian widths corresponding to the telescope resolution. Top right, bottom left, bottom right: Images of the diffuse emission in different energy intervals. After removing the compact sources, whose positions are marked by crosses, the image was smoothed by a gaussian of variable width (full width half maximum=12'–40') to obtain a uniform r.m.s. deviation of the

brightness over the image. The green threshold, then blue contours, correspond to 2σ , 3σ , 4σ , ... The 1σ deviation corresponds to (top right) $1.3 \times 10^{-7} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$ in the 5–8.5 keV band (which comprises 95% of the 6.7 keV emission), (bottom left) $1.4 \times 10^{-7} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$ in the 8.5–11 keV band, and (bottom right) $1.9 \times 10^{-7} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$ in the 12–17 keV band. In this bottom right figure, contours of the ^{13}CO radio emission from molecular gas¹³ are also shown in red, and the prominent gas clouds are labelled.

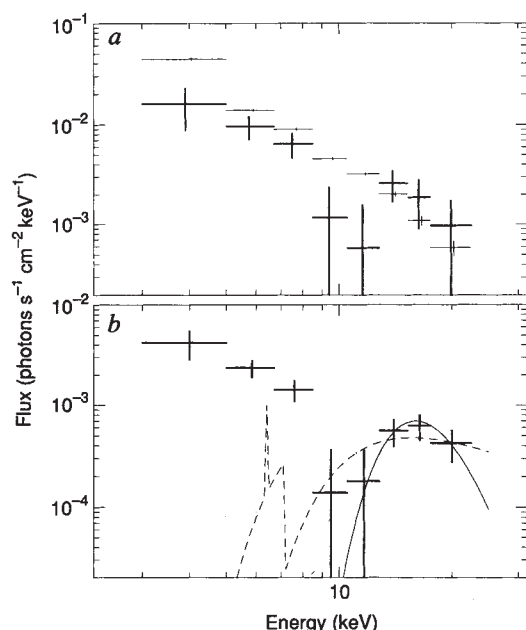


FIG. 2 *a*, Bold crosses: a spectrum of the diffuse emission from the $1^\circ \times 1.5^\circ$ ellipse (see Fig. 1, top right) with projected point sources excluded. Thin crosses: the sum of the three excluded sources. The energy bins are wider than the telescope resolution. A strong absorption at 8–11 keV features in the diffuse spectrum but not in the spectrum of the point sources. *b*, Crosses: the spectrum of the region 30' in diameter (see Fig. 1, top right). Solid line: best fit to the five harder energy bins — bremsstrahlung with $T_e = 2.8$ keV, $N_H = 9 \times 10^{24}$ cm $^{-2}$. Dashed line: a 90% limit model to the same five bins (see Fig. 3) — power law with the photon index -1.5 and $N_H = 1.8 \times 10^{24}$ cm $^{-2}$. The whole spectrum cannot be accounted for by a single model.

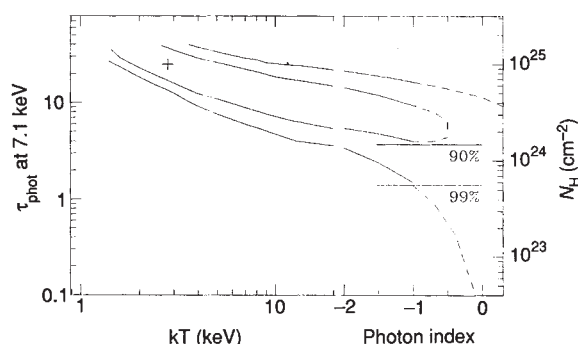


FIG. 3 Contours of constant χ^2 for fitting the harder five energy bins of the spectrum shown in Fig. 2*b* (see text). Axes are absorption column versus steepness of the incident spectrum assumed in the model. The bremsstrahlung and the power law models give similar $\chi^2_{\min}$ and best fit values of N_H . We therefore chose to show the contours for the bremsstrahlung model (N_H versus T_e), and to append those for the power law model (N_H versus the photon index) for spectra harder than bremsstrahlung. A cross marks $\chi^2_{\min} = 0.3$, with (5–3) degrees of freedom. The contours are $\chi^2_{\min} + 2.71$ and $\chi^2_{\min} + 6.63$, so that to determine the 90% and 99% lower limits on absorption column we should take the lowest point of the first and the second contour, respectively¹¹. Practically, this is equivalent to assuming the most conservative (flat) spectrum and determining the lower limit on the remaining parameter, N_H . Contours are dashed where the incident energy spectra are of increasing (positive) index.

example, there may indeed be a plasma, with a moderate temperature of 1–3 keV, located outside the opaque clouds (and therefore having much less absorption in the spectrum), with the density and the mass negligible with respect to those of the molecular gas. Contrary to our previous estimate³, the 6.4 keV neutral iron fluorescent line, another feature of the scattered spectrum, should be much weaker than the observed 6.7 keV line because there is such a high intrinsic absorption in the scattering gas.

We conclude that the observed diffuse emission above 10 keV is most probably due to scattering of incident X-rays occurring in the dense massive clouds of molecular gas that reside in the region of Galactic Centre. The lower-energy portion of the scattered emission does not escape the molecular gas because of photoabsorption. This explanation is supported by both the spatial similarity of the harder X-ray emission and the molecular gas distribution, and the strong absorption observed in its spectrum at 8–11 keV. The lower-energy emission and the 6.7-keV line, found in previous observations, may be explained by a hot plasma as before. The temperature of this plasma, however, need not be so improbably high as previously calculated. $\square$

Received 22 February 1993; accepted 18 May 1993.

1. Yamauchi, S., Kawada, M., Koyama, K., Kunieda, H. & Tawara, Y. *Astrophys. J.* **365**, 532–538 (1990).
2. Townes, C. H. in *IAU Symp. 136: The Centre of the Galaxy* (ed. Morris, M.) 1–20 (Kluwer, Dordrecht, 1989).
3. Sunyaev, R. A., Markevitch, M., & Pavlinsky, M. *Astrophys. J.* **407**, 606–610 (1993).
4. Watson, M. G., Willingale, R., Grindlay, J. E. & Hertz, P. *Astrophys. J.* **250**, 142–154 (1981).
5. Skinner, G. K., Willmore, A. P., Eyles, C. J. *et al.* *Nature* **330**, 544–547 (1987).
6. Nottingham, M. R., Skinner, G. K., Willmore, A. P. *et al.* *Astr. Astrophys. Suppl.* **97**, 165–167 (1993).
7. Vainshtein, L. A., & Sunyaev, R. A. *Soviet Astr. Lett.* **6**, 353–356 (1980).
8. Pavlinsky, M. N., Grebenev, S. A. & Sunyaev, R. A. *Soviet Astr. Lett.* **18**, 116–120 (1992).
9. Güsten, R. in *IAU Symp. 136: The Centre of the Galaxy* (ed. Morris, M.) 89–106 (Kluwer, Dordrecht, 1989).
10. Berestetskii, V. B., Lifshitz, E. M., & Pitaevskii, L. P. *Quantum Electrodynamics* (Nauka, Moscow, 1989).
11. Bevington, P. R., & Robinson, D. K. *Data Reduction and Error Analysis for the Physical Sciences* (McGraw-Hill, New York, 1992).
12. Kawai, N. *et al.* *Astrophys. J.* **330**, 130–141 (1988).
13. Heiligman, G. M. *Astrophys. J.* **314**, 747–765 (1987).
14. Fabian, A. C. *Nature* **269**, 672–673 (1977).

ACKNOWLEDGEMENTS. We thank E. Churazov for discussion, and all the members of the GRANAT team for support.

A molecular photoionic AND gate based on fluorescent signalling

A. Prasanna de Silva, H. Q. Nimal Gunaratne & Colin P. McCoy

School of Chemistry, Queen's University, Belfast BT9 5AG, UK

MOLECULES that perform logic operations are prerequisites for molecular information processing and computation^{1–11}. We^{12,13} and others^{14–16} have previously reported receptor molecules that can be considered to perform simple logic operations by coupling ionic bonding or more complex molecular-recognition processes with photonic (fluorescence) signals: in these systems, chemical binding (the 'input') results in a change in fluorescence intensity (the 'output') from the receptor. Here we describe a receptor (molecule (1) in Fig. 1) that operates as a logic device with two input channels: the fluorescence signal depends on whether the molecule binds hydrogen ions, sodium ions or both. The input/output characteristics of this molecular device correspond to those of an AND gate.

In our earlier studies^{12,13} we were concerned with exploiting photoinduced electron transfer (PET) to develop chemical sensors¹⁷. Molecule (2)¹² (Fig. 1) produced a fluorescence signal only when the concentrations of protons was low (about 10^{-6} M). When the proton concentration signal rose to 10^{-2} M, the

fluorescence signal fell to nearly zero. According to the perspective of molecular devices, this can be regarded as a NOT function. Molecule (3)¹³ gave a fluorescence signal only when the proton 'input' was high (about 10^{-3} M), corresponding to pass (YES) logic function. Pass (YES) logic devices will be essential for photoionic signal processing, as they interconvert the two different kinds of signal (ionic to photonic).

These devices have one-input/one-output functionality. Complex logic operations will also need two-input devices such as AND and OR gates. Although the receptors in refs 14–16 are arguably such devices, they do not possess a full truth table as the unprotonated forms simply will not bind the substrate. Our receptor (1) is an anthracene derivative; such compounds have previously been shown to undergo PET processes from tertiary amines and benzocrown ethers separately. These transfer processes could be suppressed by injection of protons¹³ and sodium ions¹⁸, respectively, to cause 'switching on' of fluorescence. Therefore molecule (1), based on the 'fluorophore-spacer₁-receptor₁-spacer₂-receptor₂' format could be

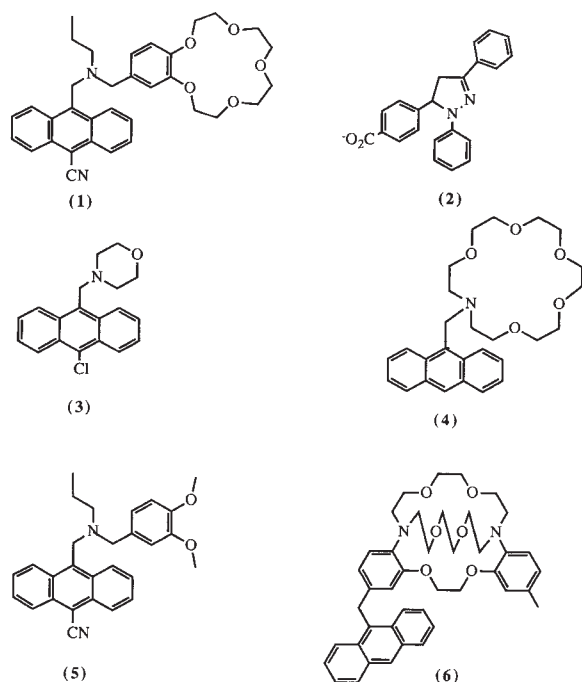


FIG. 1 Materials for molecular photoionic logic devices. Synthesis of molecule (1): benzo-15-crown-5-ether-aldehyde²⁸ was condensed with 1-propylamine, reduced with sodium borohydride and finally reacted with 10-cyano anthracene-9-yl-methyl bromide²⁹. Elemental analysis for (1): C, 73.5; H, 6.6; N, 5.0%. $C_{34}H_{38}N_2O_5$ requires: C, 73.6; H, 6.9; N, 5.1%. Spectrometric parameters: 1H NMR ($CDCl_3$): δ = 7.57–8.61 p.p.m. (8H, multiplet (m), anthracene hydrogen (Anth-H) 6.66–6.70 (3H, m, aromatic hydrogen (Ar-H), 4.54 (2H, singlet s, Anth-CH₂), 3.75–4.09 (16H, m, OCH₂), 3.49 (2H, s, ArCH₂), 2.49 (2H, triplet t, NCH₂CH₂), 1.59 (2H, m, CH₂CH₃), 0.73 (3H, t, CH₂CH₃). Mass/charge ratio m/z = 554 (M^+). Infrared spectroscopy ν_{max} (KBr) = 2,230 cm^{-1} (C=N bond). Ultraviolet λ_{max} (ϵ_{max}) (CH_3OH): 410 nm (8,300 $M^{-1} cm^{-1}$), 389 (9,700), 369 (7,700). Molecule (5) similarly synthesized with 3,4-dimethoxy-benzaldehyde as a starting material. Elemental analysis for (5): C, 79.2; H, 6.6; N, 6.6%. $C_{28}H_{28}N_2O_2$ requires: C, 79.2; H, 6.7; N, 6.6. Spectrometric parameters: 1H NMR ($CDCl_3$): δ = 7.57–8.60 (8H, m, Anth-H), 6.69 (3H, brs, Ar-H), 4.55 (2H, s, Anth-CH₂), 3.82 and 3.71 (3H and 3H, s and s, 4'-OCH₃ and 3'-OCH₃), 3.51 (2H, s, Ar-CH₂), 2.49 (2H, t, NCH₂CH₂), 1.59 (2H, m, CH₂CH₃), 0.75 (3H, t, CH₂CH₃). m/z = 424 (M^+). Infrared ν_{max} (KBr): 2,206 cm^{-1} (C=N bond). Ultraviolet λ_{max} (ϵ_{max}) (CH_3OH): 410 nm (8,300 $M^{-1} cm^{-1}$), 389 (9,700), 370 (7,800). Molecules (2), (3), (4) and (6) are described in refs 12, 13, 27 and 30 respectively.

TABLE 1 Fluorescence output for different input combinations

Molecule	Ionic input 1 (concentration)	Ionic input 2 (concentration)	Fluorescence output quantum yield	Fluorescence enhancement factor
(1)	0	0	0.011(5)	1.0 (standard)
(1)	H ⁺ (10^{-3} M)	0	0.020	1.7
(1)	0	Na ⁺ (10^{-2} M)	0.013	1.1
(1)	H ⁺ (10^{-3} M)	Na ⁺ (10^{-2} M)	0.068	6.0
(2)	H ⁺ (10^{-6} M)	—	0.13	1.0 (standard)
(2)	H ⁺ (10^{-2} M)	—	0.003	0.023
(3)	H ⁺ (10^{-8} M)	—	0.003	1.0 (standard)
(3)	H ⁺ (10^{-3} M)	—	0.42	140
(4)	0	0	0.003	1.0 (standard)
(4)	Na ⁺ (10^{-2} M)	0	0.053	18
(4)	0	K ⁺ (10^{-2} M)	0.14	47
(4)	Na ⁺ (10^{-2} M)	K ⁺ (10^{-2} M)	0.14	46
(5)	0	0	0.004	1.0 (standard)
(5)	H ⁺ (10^{-2} M)	0	0.008	2.0
(5)	0	Na ⁺ (10^{-2} M)	0.004	1.0
(5)	H ⁺ (10^{-2} M)	Na ⁺ (10^{-2} M)	0.008	2.0

Experimental conditions: 10^{-5} M (1) or (5) dissolved in mixture of CH_3OH and 2- C_3H_7OH (1:1 v/v) with $CH_3SO_3^-$ counterions, excited at 387 nm, fluorescence emission maximum was at 446 nm. 10^{-5} M (2) in $CH_3OH:H_2O$ (1:4, v/v) with Cl^- counter ions, excited at 359 nm; emission maximum was at 475 nm. 10^{-6} M (3) in $CH_3OH:H_2O$ (1:4 v/v) excited at 377 nm; emission maximum was at 428 nm. 10^{-5} M (4) in CH_3OH with $CH_3CO_2^-$ counter ions, excited at 368 nm; emission maximum was at 414 nm.

expected to fluoresce only when the tertiary amine and the benzocrown ether units are receiving proton and sodium ion input, respectively, at sufficiently high concentrations (Fig. 2).

The successful operation of molecule (1) as a two-input AND logic gate depends on the fact that tertiary amines and benzocrown ethers are highly selective for protons¹⁹ and for sodium ions²¹, respectively, so each receptor will accept its designated ion while largely ignoring the other ion(s) present. This allows the molecule to self-select its input ionic signals into the appropriate site once it has been provided with a pool of protons and sodium ions (at concentrations chosen by the operator). No 'wiring' is needed for the molecular device to process the ionic input signals and produce fluorescence output. Fluorescence signals from single molecules can be received and analysed by spectroscopy^{21,22}. In contrast, molecular wires²³ or scanning probe tips^{24–26} are necessary to operate all-electronic molecular devices, even with a single input.

Our results (Fig. 3, Table 1) show that these expectations are largely realized. The fluorescence output of molecule (1) increases only by a factor of 1.1 when 10^{-2} M sodium ions alone are present and by a factor of 1.7 when 10^{-2} M protons alone are present. In contrast, when both inputs are simultaneously applied, the fluorescence output rises by a factor of 6.0. Thus, the output is 'high' only when both sodium ions and protons are coincident on (1), and remains 'low' under the other three input signal conditions. The increased output in the presence of protons alone probably arises because the oxidation potential of the remote benzocrown ether increases on protonation of its aminomethyl side chain, reducing the PET rate. The alternative possibility, that solvated protons are complexed by the benzocrown ether²⁰, seems less likely, as the model molecule (5) devoid of a crown ether ring shows a fluorescence output enhancement of 2.0 both with protons alone and when the protons are joined by sodium ions.

Detailed examination of the fluorescence quantum yield of molecule (1) as a function of the concentrations of sodium ions

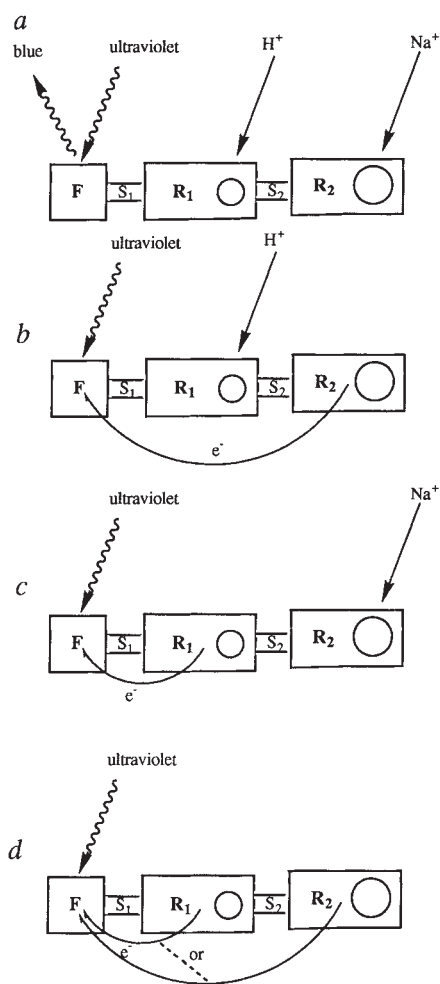


FIG. 2 Design principle of the molecular photoionic AND gate (1). F, fluorophore; S_1 , S_2 , spacers; R_1 , proton receptor; R_2 , sodium ion receptor. The electron transfer is schematic only as the process may occur through bonds or through space.

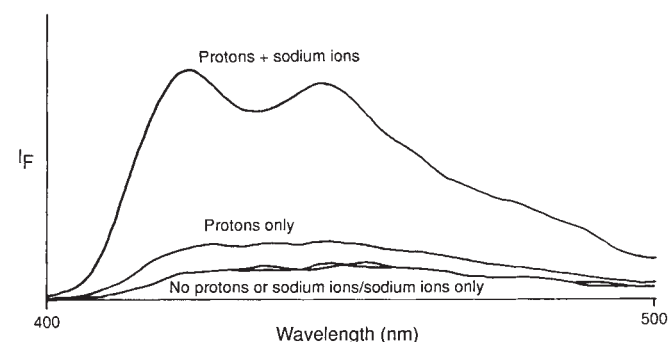


FIG. 3 Fluorescence emission spectroscopy of molecule (1) for different ionic inputs. Detailed conditions are given in Table 1. The fluorescence emission spectra (and the associated absorption spectra) hardly change position or shape. The dominant variable by far is the fluorescence quantum yield.

and protons allows us to calculate the binding constants (K_b , in M^{-1} , given in log units) of the interactions between the benzocrown ether and sodium ion, and the tertiary amine and proton. In the presence of limiting high concentrations of the other ion, these are 2.6 and 4.5, respectively. As $K_b = [ML]/[M][L]$ (where the concentrations refer to the metal–ligand complex, free metal and free ligand respectively), switching the fluorescence output from ‘low’ ($[ML] \ll [L]$) to ‘high’ ($[ML] \gg [L]$) requires the concentration of metal ions, $[M]$, considerably to exceed K_b^{-1} . The fluorescence output will revert to ‘low’ when the ionic signal is reduced to well below K_b . The switching is reversible, as demonstrated by the quenching of the fluorescence output to baseline levels by adding morpholine or 18-crown-6 ether which scavenge protons and sodium ions respectively.

A general photoionic logic system also requires devices with the OR function. An approximate photoionic OR gate would be molecule (4), which has poor chemoselectivity. Either sodium or potassium ions in the ion pool can bind to the azacrown ether receptor of (4) causing the fluorescence output signal to ‘switch on’. However, the magnitude of the ‘high’ fluorescence output signal depends somewhat on the nature of the ionic input signal (Table 1). A better OR function can be found in the cryptand (6)²⁷ where 5×10^{-2} M potassium or rubidium ions enhance the fluorescence by factors of 9.4 and 10.6 in CH_3OH .

Although molecules (1)–(4) provide a rudimentary foundation for future molecular photoionic devices, they have other immediate uses. Molecules (2) and (3) can allow real-time fluorescence monitoring of ion concentrations and their distribution, and molecules such as (1) can allow direct fluorescence monitoring of situations where two (or more) chemical species must come together to cause a vital process. □

Received 8 April; accepted 10 May 1993.

- Carter, F. L., Siatkowski, R. E. & Wohltjen, H. (eds) *Molecular Electronic Devices* (Elsevier, Amsterdam, 1988).
- Lehn, J.-M. *Angew. Chem. int. Ed. Engl.* **27**, 89–112 (1988).
- Lehn, J.-M. *Angew. Chem. int. Ed. Engl.* **29**, 1304–1319 (1990).
- Anelli, P. L. et al. *J. Am. chem. Soc.* **114**, 193–218 (1992).
- Bryce, M. et al. *Chem. Brit.* **27**, 707–731 (1991).
- Chambron, J. C., Heitz, V. & Sauvage, J.-P. *J. chem. Soc. chem. Commun.*, 1131–1133 (1992).
- Wasielewski, M. R., O’Neil, M. P., Gosztola, D., Niemczyk, M. P. & Svec, W. A. *Pure appl. Chem.* **64**, 1319–1325 (1992).
- Wild, U. P., Bernet, S., Kohler, B. & Renn, A. *Pure appl. Chem.* **64**, 1335–1342 (1992).
- Simon, J., Engel, M. K. & Soulie, C. *New J. Chem.* **16**, 287–293 (1992).
- Ashton, P. R., Philp, D., Spencer, N. & Stoddart, J. F. *J. Chem. Soc. chem. Commun.*, 1124–1128 (1992).
- Ashton, P. R., Johnston, M. R., Stoddart, J. F., Tolley, M. S. & Wheeler, J. W. *J. Chem. Soc. chem. Commun.*, 1128–1131 (1992).
- de Silva, A. P., de Silva, S. A., Dissanayake, A. S. & Sandanayake, K. R. A. S. *J. Chem. Soc. chem. Commun.*, 1056–1058 (1989).
- de Silva, A. P. & Rupasinghe, R. A. D. D. *J. chem. Soc. chem. Commun.*, 1669–1670 (1985).
- Huston, M. E., Akkaya, E. U. & Czarnik, A. W. *J. Am. chem. Soc.* **111**, 8735–8737 (1989).
- Hosseini, M. W., Blacker, A. J. & Lehn, J.-M. *J. Am. chem. Soc.* **112**, 3896–3904 (1990).
- Van Arman, S. A. & Czarnik, A. W. *Supramolec. Chem.* **1**, 99–101 (1993).
- Bissell, R. A. et al. *Chem. Soc. Rev.* **21**, 187–195 (1992).
- de Silva, A. P. & Sandanayake, K. R. A. S. *J. chem. Soc. chem. Commun.*, 1183–1185 (1989).
- Smith, J. W. in *The Chemistry of the Amino Group* (ed. Patai, S.) Ch. 4 (Interscience, London, 1968).
- Izatt, R. M., Pawlak, K., Bradshaw, J. S. & Bruening, R. L. *Chem. Rev.* **91**, 1721–2085 (1991).
- Orrit, M. & Bernard, J. *Phys. Rev. Lett.* **65**, 2716–2719 (1990).
- Basche, Th. & Moerner, W. E. *Nature* **355**, 335–337 (1992).
- Kugimiya, S. I., Lazrak, T., Blanchard-Desce, M. & Lehn, J.-M. *J. chem. Soc. chem. Commun.*, 1179–1182 (1991).
- Binnig, G., Rohrer, H., Gerber, C. & Weibel, E. *Appl. Phys. Lett.* **40**, 178–180 (1982).
- Binnig, G. & Rohrer, H. *Angew. Chem. Int. Edn. Engl.* **26**, 606–614 (1987).
- Frommer, J. *Angew. Chem. Int. Edn. Engl.* **31**, 1289–1328 (1992).
- de Silva, A. P., Gunaratne, H. Q. N. & Sandanayake, K. R. A. S. *Tetrahedron Lett.* **31**, 5193–5196 (1990).
- Hyde, E. M., Shaw, B. L. & Shepherd, I. J. *J. chem. Soc. Dalton Trans.*, 1696–1705 (1978).
- Perera, K. S. D., thesis, Queen’s Univ. of Belfast (1989).
- de Silva, A. P. & de Silva, S. A. *J. chem. Soc. Commun.*, 1709–1710 (1986).

ACKNOWLEDGEMENTS. We thank A. G. Smith and I. Campbell for assistance, and the SERC and DENI for support.

Anomalous CH_4 and NH_4^+ concentrations at an unsedimented mid-ocean-ridge hydrothermal system

M. D. Lilley*, D. A. Butterfield†, E. J. Olson*, J. E. Lupton‡, S. A. Macko§ & R. E. McDuff*

* School of Oceanography WB-10, University of Washington, Seattle, Washington 98195, USA

† NOAA, Pacific Marine Environmental Laboratory, 7600 Sand Point Way, Seattle, Washington 98115, USA

‡ NOAA, Pacific Marine Environmental Laboratory, Hatfield Marine Science Center, Newport, Oregon 97365, USA

§ Department of Environmental Sciences, University of Virginia, Charlottesville, Virginia 22903, USA

SINCE the discovery in 1977 of sea-floor hydrothermal systems, the study of the chemistry of the venting fluids has transformed our understanding of the geochemical cycles that influence the composition of sea water and the ocean crust. With few exceptions (Guaymas basin being the most notable), the vent systems studied so far are free of sedimentary influence and the chemistry of the fluids can be explained on the basis of interactions between sea water and basalt. Such fluids typically contain low methane concentrations, ranging from 50 to 120 μM (refs 1–7), and ammonium concentrations less than 10 μM (ref. 8). Here we report CH_4 and NH_4^+ concentrations from the Endeavour segment of the Juan de Fuca Ridge which are many times greater than those measured previously at any unsedimented mid-ocean ridge. The $^{13}\text{C}/^{12}\text{C}$ ratio of this CH_4 is the lowest yet found in any hydrothermal environment, implying an unusual source. We attribute these high CH_4 and NH_4^+ concentrations to the decomposition of sub-sea-floor organic matter associated with sediments buried at an earlier stage of the ridge's evolution. These data illustrate that the organic geochemistry of unsedimented ridges may be more complex than suspected hitherto.

Located at a depth of 2,200 m on the Endeavour segment of the Juan de Fuca Ridge, the main Endeavour hydrothermal field ($47^\circ 57' \text{N}$, $129^\circ 06' \text{W}$) is characterized by steep-sided sulphide structures up to 30 m across at their base and up to 20 m high⁹. Most of these structures (Fig. 1) discharge hot fluids from multiple black smoker chimneys and from numerous overhanging ledges ('flanges') where ponded fluids overflow the outer edges. The chemistry of this vent field is variable, as the CH_4 and NH_4^+ data show (Fig. 2). These data show a systematic trend with position in the vent field (compare Figs 1 and 2), which corresponds to a north–south thermal gradient in fluid temperature of 317–368 $^\circ\text{C}$. These data are compared with other hydrothermal systems in Table 1. With the notable exception of the system at 9°N on the East Pacific Rise¹⁰, the Endeavour vent field also contains the lowest measured Cl^- concentrations (20% of that in sea water) in any reported vent fluids. Although small reductions in Cl^- concentration from the parent sea water may be attributed to precipitation of a chloride-bearing mineral¹¹, the large reductions seen at Endeavour can only be the result of phase separation and segregation of low- Cl^- vapour from high- Cl^- brine¹² below the sea floor.

The variability of CH_4 with Cl^- (Fig. 2b) is, however, too large to have been generated solely by supercritical phase separation, where a small volume of high-salinity brine separates from the remaining vapour phase¹³. If all the CH_4 remained in the vapour phase, the small volume reduction accompanying this process would increase the CH_4 concentration by a few per cent at most. Rather, if we assume initially high

concentrations in the original fluid, the two trends of CH_4 with Cl^- , representing the northern and southern ends of the vent field, are consistent with a subcritical boiling model and variable extents of boiling¹⁴. This is in accord with the thermal structure of the vent field: fluids from the southern group of structures are 10–38 $^\circ\text{C}$ hotter than the northern set and would encounter the two-phase boundary at greater depth. This argument does not rule out the possibility that the fluids have undergone supercritical phase separation deeper within the system prior to the subcritical boiling as the fluid follows the boiling curve to the sea floor.

Thus at Endeavour, the phase separation that seems to have occurred will have brought about a further increase of the CH_4 and NH_4^+ concentrations in the low- Cl^- fluids, but it cannot be the primary cause of the anomalously high concentrations. To achieve the high measured CH_4 and NH_4^+ by simply partitioning gas from a typical mid-ocean-ridge fluid (with CH_4 and NH_4^+ concentrations much lower than Endeavour) into the vapour phase would require even steeper trends of CH_4 and NH_4^+ with Cl^- than seen in Fig. 2. The $\text{CH}_4/{}^3\text{He}$ ratios (Table 1) strengthen this point. As the solubilities of these gases are similar, they would have similar concentrations in the vapour phase, and an additional source of CH_4 is required to produce the high ratio found. More evidence comes from data obtained at Axial seamount and the 9°N site on the East Pacific Rise. Both sites show phase separation^{10,14} yet the fluids contain the normal low concentrations of CH_4 and NH_4^+ associated with unsedimented ridges (M.D.L., unpublished data).

The potential sources of CH_4 in hydrothermal systems include the thermal breakdown of organic material, bacterial production, outgassing of juvenile CH_4 and inorganic synthesis¹⁵. Methane isotopic signatures and concentrations found in basalts suggest the fairly low concentration of CH_4 found at 21°N is juvenile CH_4 extracted from the rock¹, whereas the high

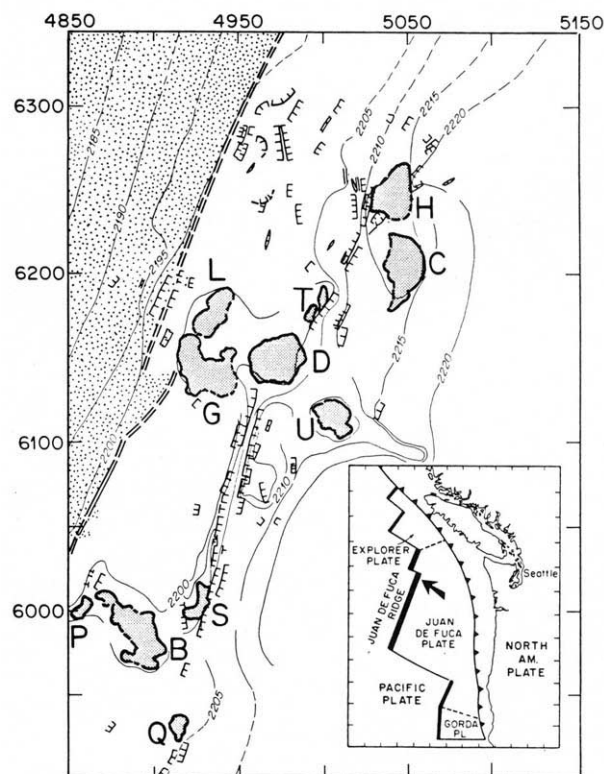


FIG. 1 Map showing relative locations of sulphide structures at the main vent field, Endeavour segment of Juan de Fuca Ridge. Axes are in metres. Map adapted from ref. 9.

TABLE 1 Comparison of endmember hydrothermal fluids

Component	Endeavour	SJdF ³	21° N EPR ^{1,2,8}	11° N EPR ^{4,5}	13° N EPR ⁴⁻⁶	MAR ⁷	Okinawa ¹⁷	Guaymas ^{16,32}
CH ₄ (mmol kg ⁻¹)	1.8–3.4	0.08–0.12	0.06–0.09	0.067–0.116	0.051	0.062	2.4	2.0–6.8
NH ₄ ⁺ (mM)	0.64–0.95	< 0.01*	< 0.01				5.0	10.3–15.6
CO ₂ (mmol kg ⁻¹)	11.6–18.2	3.7–4.5	5.7		11.8–18.4		198	
H ₂ (mmol kg ⁻¹)	0.16–0.42	0.20–0.53	0.23–1.70	0.47	0.14			
Rn (d.p.m. kg ⁻¹)	100–1,800		321					
CH ₄ / ³ He (× 10 ⁻⁶)	107–217		2.5	5.3	3.5	2.6	324	3,100
δ ¹³ CH ₄	–48.4 to –55.0	–17.8 to –20.8	–16		–16.6 to –19.5		–36.1 to –40.7	–43.2 to –50.8
CH ₄ /C ₂₊	449–539	1,600–5,200	303–1428					78–690

SJdF, south Juan de Fuca Ridge; EPR, East Pacific Rise; MAR, Mid-Atlantic Ridge. For the Endeavour gas samples, we compared the results of 'gas-tight' samplers directly with those of traditional 'major' samplers as were used to collect the SJdF, EPR and Guaymas samples. The ratio of concentration obtained by gas-tight/major samplers was 2–3, and this should be borne in mind when comparing the data. It is likely that this ratio could be higher in a system with higher total gas contact, such as Guaymas. Data reference for Endeavour is this work.

* M.D.L., unpublished data; other unsedimented sites where NH₄⁺ concentrations are <0.01 mM include Axial seamount, Juan de Fuca Ridge and 9° N EPR data (M.D.L., unpublished); all other references indicated by superscript.

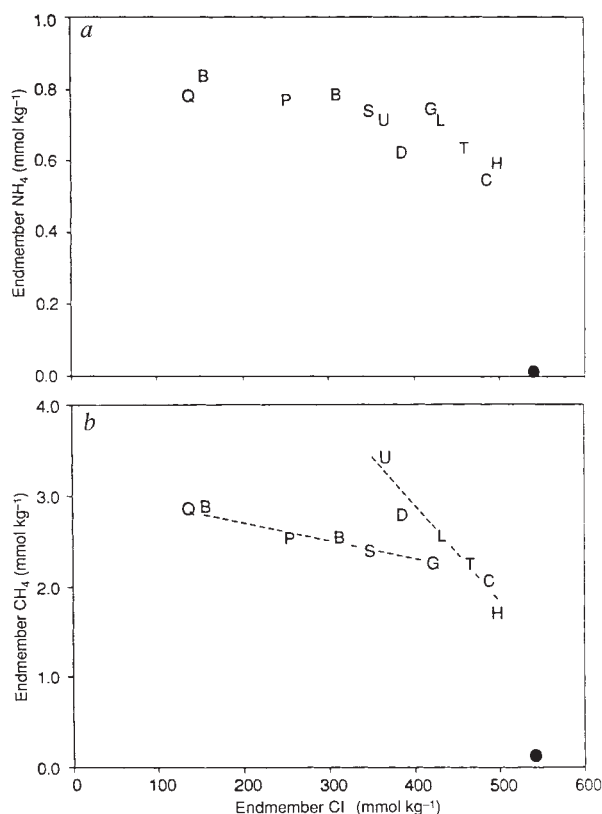
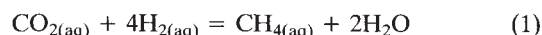


FIG. 2a, Plot of endmember NH₄⁺ against endmember Cl⁻. The letter designations used for the data points refer to the individual sulphide structures depicted in Fig. 1 and (●) represents the highest concentration typically found in unsedimented ridge systems (Tables 1 and refs therein). Endmember refers to the concentration of a chemical species in pure hydrothermal fluid. As the sampling process usually entrains some sea water and hydrothermal fluids are known to be devoid of Mg, measured concentrations are extrapolated to zero Mg to obtain endmember values. b, Plot of endmember CH₄ against endmember Cl⁻. Fluid samples for volatile analyses were collected by the submersible Alvin in gas-tight titanium samplers. The gases were vacuum-extracted from the fluids at sea and sealed in Pyrex break seals for subsequent analysis by shore-based gas chromatography.

concentrations from Guaymas Basin are primarily derived from thermogenic decomposition of overlying sediments¹⁶. The CH₄ isotopic signature of Endeavour fluids (Fig. 3, Table 1) is significantly lighter than at any other unsedimented ridge and is even lighter than fluids from systems with known sedimentary influences (such as Guaymas basin¹⁶ and Okinawa basin¹⁷). Additionally, the CH₄/³He ratio is very high and the CH₄/C₂₊ ratio low for a sediment-free system (Table 1). On the basis of these data and our current understanding of the origins of CH₄ in hydrothermal systems¹⁵, we believe that thermal decomposition of organic material is the most likely source. The possibility of microbial CH₄ production in the downwelling portion of the hydrothermal convection cell cannot, however, be ruled out, as CH₄ with a δ¹³C of –55‰ (PDB) is within the range of bacterial production¹⁸. Additionally, the isotopic fractionation associated with microbial reduction of CO₂ to CH₄ under the variable temperatures and pressures of hydrothermal systems has not been documented².

Using the data from Table 1, the equilibrium constant (log *K* = 8.45 at 350°C, 22 MPa) derived from SUPCRIT92¹⁹ for the reaction



and the thermodynamic data cited in ref. 3, we calculate that the CH₄ concentrations at Endeavour are about two orders of magnitude too high to be in equilibrium with the measured H₂ and CO₂ levels under these conditions.

Ammonium (NH₄⁺) in hydrothermal systems could come from extraction from basalt, high-temperature N₂ fixation, and thermal or microbial degradation of organic matter. Its concentration does not exceed 10 μM at any other sediment-starved mid-ocean ridge (Table 1), and it is unlikely that processes such as N₂ fixation and extraction from basalt would occur only at Endeavour. Indeed, N₂ fixation can be ruled out at Endeavour as no dissolved N₂ deficit was found in the fluids. Thus, we conclude that the NH₄⁺ is produced by the decomposition of organic material.

One possible organic source is the diverse animal community, including microbial mats, covering the sulphide structures. Both thoroughly and incompletely pyritized worm tubes are commonly observed embedded in the Endeavour sulphides, suggesting that high-temperature fluids decompose animal tissue which has been incorporated into the sulphide structure. However, our nitrogen isotope data on vent animals and fluids do not support

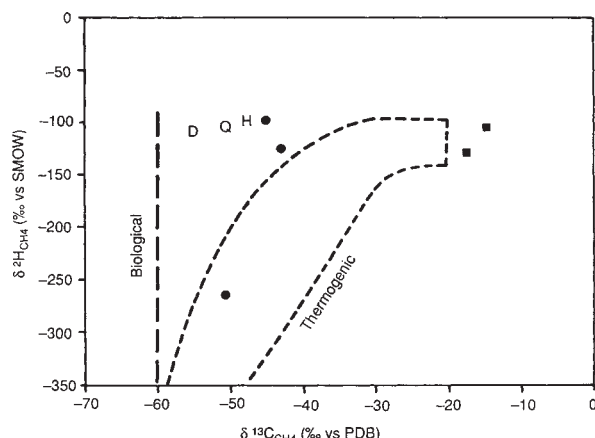


FIG. 3 Isotopic composition of CH_4 from Endeavour (H, Q and D, as in Fig. 1), 21°N EPR¹ (■), and Guaymas basin¹⁴ (●). The data fields marked 'Biological' and 'Thermogenic' are as presented in ref. 15. Additional $\delta^{13}\text{CCH}_4$ data for which the corresponding hydrogen isotope data are not available are listed in Table 1.

the hypothesis that this is the main source of NH_4^+ . The average $\delta^{15}\text{N}$ value of vent animals (tube worms, palm worms, limpets) is -1.4‰ , whereas the $\delta^{15}\text{N}\text{‰}$ for the vent fluid NH_4^+ averages (relative to air) $+12.4\text{‰}$. The isotopically light CH_4 (-55‰ relative to PDB) implies that decomposition of the organic material is in an early stage^{20,21} and has not proceeded far enough to produce a reservoir. Likewise, the NH_4^+ would probably be as much or slightly more depleted in the heavier isotope than its source material and, therefore, isotopically lighter than the source animals rather than heavier^{22,23}. The $\delta^{15}\text{NH}_4^+$ values found here are similar to values reported for $\delta^{15}\text{NH}_4^+$ in sedimentary pore waters²⁴.

An alternative source of organic material could be sediments similar to the Pleistocene turbidite deposits that buried the Middle Valley section of the Juan de Fuca Ridge about 40 km north of the Endeavour site. The Endeavour vent field rests on a recent, elongate volcanic edifice which stands 200 m shallower than these turbidite flows. It is possible that these turbidite deposits coincided with a waning magmatic period when this section of ridge was deeper^{25,26}, and that subsequent lava flows built the volcanic edifice on buried sediments. The possibility of surface-ponded sediment influence on the downwelling portion of the convection cell should also be considered but is believed to be unlikely as the nearest important sediment ponds, both on- and off-axis, are 10 km away (P. Johnson, personal communication). This distance is much larger than the expected extent of hydrothermal convection cells²⁷. Whatever the source, elevated concentrations of Rb, Cs, I and B, a B isotopic anomaly (D.B., unpublished data), high ^{222}Rn values (ref. 28 and Table 1) and relatively low Eu concentrations (G. Klinkhammer, personal communication) are consistent with a sedimentary influence on the fluids venting along the Endeavour segment.

Regardless of the source of organic material, a simple mass balance implies that the current output of CH_4 at Endeavour cannot be in steady state. If CH_4 is produced from organic carbon of Redfield composition²⁹ $(\text{CH}_2\text{O})_{106}(\text{NH}_3)_{16}(\text{H}_3\text{PO}_4) = 53\text{CO}_2 + 53\text{CH}_4 + 16\text{NH}_3 + \text{H}_3\text{PO}_4$, and we assume $1.4 \times 10^8 \text{ mol CH}_4 \text{ yr}^{-1}$ as the flux³⁰, then 3,300 metric tonnes yr^{-1} of organic material would be required to maintain the CH_4 flux. Put another way, if the areal extent of a subsurface sediment layer influencing this hydrothermal field were 1 km^2 , and the sediment has a bulk density of 2 g cm^{-3} with an organic carbon concentration of 0.5%, an equivalent sediment thickness of $\sim 0.3 \text{ m yr}^{-1}$ would be required to maintain the CH_4 flux. Similar results follow if the organic material is assumed to be kerogen ($\text{C}_{10}\text{H}_{14}\text{O}_4\text{N}$) and decomposition produces CO_2 and CH_4 in roughly equal amounts.

Radiometric dating of active chimneys from the Endeavour vent field yields ages greater than 200 years in some cases³¹ and implies that the overall structures are much older still. A system age of this magnitude, coupled with the light $\delta^{13}\text{C}$ of the CH_4 , requires either that a very thick sediment layer is available to this system or that the hydrothermal convection cell has encountered the sediment layer fairly recently, implying that the high concentrations of CH_4 and NH_4^+ are transient. We believe that the latter is the most plausible explanation of our data. The Endeavour vent field represents a new type of ridge-crest hydrothermal system in that it not only undergoes phase separation but also, although it appears sediment-free, shows many of the chemical characteristics of sediment-covered and back-arc systems. Planned analyses for the cosmogenic isotope ^{10}Be on Endeavour samples should indicate more clearly the importance of sediments relative to other potential sources of organic material in this system. □

Received 13 January; accepted 27 April 1993.

1. Welhan, J. A. & Craig, H. in *Hydrothermal Processes at Seafloor Spreading Centers* (eds Rona, P. A., Bostrom, K., Laubier, L. & Smith, K. L.) 391–409 (Plenum, New York, 1983).
2. Lilley, M. D., Baross, J. A. & Gordon, L. I. in *Hydrothermal Processes at Seafloor Spreading Centers* (eds Rona, P. A., Bostrom, K., Laubier, L. & Smith, K. L.) 411–449 (Plenum, New York, 1983).
3. Evans, W. C., White, L. D. & Rapp, J. B. *J. geophys. Res.* **93**, 15305–15313 (1988).
4. Kim, K.-R., Welhan, J. A. & Craig, H. *Eos* **65**, 973 (1984).
5. Welhan, J. A., Craig, H. & Kim, K. *Eos* **65**, 973 (1984).
6. Merlivat, L., Pineau, F. & Javoy, M. *Earth planet. Sci. Lett.* **84**, 100–108 (1987).
7. Jean-Baptiste, P., Charlou, J. L., Stievenard, M., Donval, J. P. & Mevel, C. *Earth planet. Sci. Lett.* **106**, 17–28 (1991).
8. Von Damm, K. L. *A. Rev. Earth planet. Sci.* **18**, 173–204 (1990).
9. Delaney, J. R., Robigou, V., McDuff, R. E. & Tivey, M. K. *J. geophys. Res.* (in the press).
10. Von Damm, K. L., Grebeiner, J. M. & Edmond, J. M. *Eos* **72**, 480 (1991).
11. Berndt, M. E. & Seyfried, W. E. Jr *Geochim. cosmochim. Acta* **54**, 2235–2245 (1990).
12. Butterfield, D. A. thesis, Univ. Washington (1990).
13. Bischoff, J. L. & Rosenbauer, R. J. *Am. J. Sci.* **287**, 953–978 (1987).
14. Butterfield, D. A., Massoth, G. J., McDuff, R. E., Lupton, J. E. & Lilley, M. D. *J. geophys. Res.* **95**, 12895–12921 (1990).
15. Welhan, J. A. *Chem. Geol.* **71**, 183–198 (1988).
16. Welhan, J. A. & Lupton, J. E. *Am. Assoc. Petrol. Geol. Bull.* **71**, 215–223 (1987).
17. Sakai, H. et al. *Geophys. Res. Lett.* **17**, 2133–2136 (1990).
18. Schoell, M. *Chem. Geol.* **71**, 1–10 (1988).
19. Johnson, J. W., Oelkers, E. H. & Helgeson, H. C. *Comput. Geosci.* **18**, 899–947 (1992).

20. Sackett, W. M. *Geochim. cosmochim. Acta* **42**, 571–580 (1978).
21. Rohrback, B. G., Peters, K. E. & Kaplan, I. R. *Am. Assoc. Petrol. Geol. Bull.* **68**, 961–970 (1984).
22. Macko, S. A., Engel, M. H. & Parker, P. L. in *Organic Geochemistry* (eds Engel, M. H. & Macko, S. A.) 211–224 (Plenum, New York, 1993).
23. Rohrback, B. G., Peters, K. E., Sweeney, R. E. & Kaplan, I. R. *Advances in Organic Geochemistry 1981* (ed. Bjorøy, M.) 819–823 (Wiley, New York, 1983).
24. Sweeney, R. E. & Kaplan, I. R. *Mar. Chem.* **9**, 81–94 (1980).
25. Kappel, E. S. & Ryan, W. B. F. *J. geophys. Res.* **91**, 13925–13940 (1986).
26. Tivey, M. A. & Johnson, H. P. *J. geophys. Res.* **92**, 12685–12694 (1987).
27. Cann, J. R., Strens, M. R. & Rice, A. *Earth planet. Sci. Lett.* **76**, 123–134 (1985).
28. Kadko, D. & Moore, W. *Geochim. cosmochim. Acta* **52**, 659–668 (1988).
29. Redfield, A. C., Ketchum, B. H. & Richards, F. A. in *The Sea*, Vol. II (ed. Hill, M. N.) 26–77 (Wiley, New York, 1963).
30. Rosenberg, N. D., Lupton, J. E., Kadko, D., Collier, R., Lilley, M. D. & Pak, H. *Nature* **334**, 604–607 (1988).
31. Kim, K. H. & McMurtry, G. M. *Earth planet. Sci. Lett.* **104**, 299–314 (1991).
32. Von Damm, K. L., Edmond, J. M., Measures, C. I. & Grant, B. *Geochim. cosmochim. Acta* **49**, 2221–2237 (1985).

ACKNOWLEDGEMENTS. We thank M. Dudley and E. McLaughlin for analytical assistance, the pilots, officers and crew of *Atlantis II/Alvin* for their dedication and expertise, P. Quay, D. Wilbur, R. Carmody and W. Shanks for isotopic analyses, J. Delaney, J. Baross, P. Johnson, D. Norton, K. Von Damm and J. Gieskes for discussions, K. Von Damm and M. Tivey for SUPCRIT92 data, and J. Bischoff and J. Seewald for reviews. Support was provided by the NSF.

Reduced Himalayan sediment production 8 Myr ago despite an intensified monsoon

Douglas W. Burbank*, Louis A. Derry† & Christian France-Lanord‡

* University of Southern California, Department of Geological Sciences, Los Angeles, California 90089, USA

† Hartwick College, Department of Chemistry, Oneonta, New York 13820, USA

‡ CRPG/CNRS, BP 20, 54501, Vandoeuvre-les-Nancy, France

UPLIFT of the Tibetan plateau about 7–8 Myr ago^{1,2} may have been responsible^{3–5} for the apparent intensification of the Asian monsoon^{6–10} around that time. Increases in the oceanic ⁸⁷Sr/⁸⁶Sr ratio during the Neogene period have been attributed^{11,12} to increased erosion from the Himalayan orogen. It has been suggested that the monsoonal intensification may have enhanced overall erosion rates in the region^{5,13}. If this were the case, sediment accumulation rates would have increased in the surrounding basins at this time. Here we present a reanalysis of stratigraphic data from the Indo-Gangetic foreland and the Bengal fan, which demonstrates that both of these basins experienced a decline in sediment-accumulation rates 8 Myr ago. Thus it seems that monsoonal intensification was accompanied by a decrease in mechanical weathering. This decrease could be due to reduced tectonic activity, decreased Himalayan glaciation or slope stabilization from dense plant cover.

Recent isotope studies of soil carbonates^{6–8,14} within Siwalik strata in both Pakistan and Nepal (Fig. 1) show considerable shifts in carbon and oxygen isotope ratios in the Late Miocene. These shifts have been interpreted as signs of climate change, manifested by the replacement of C3 plants (trees and shrubs) by C4 plants (grasses) and decreased soil leaching in response to increased seasonality and monsoonal development in southern Asia. Coupling of these isotope data with magnetostratigraphic studies permits the precise timing of the isotopic shift to be examined. If an intensified monsoonal climate were responsible for the shifts, one would expect generally synchronous changes across the entire sub-Himalayan region. This expectation has been supported by the observation of a coeval reduction in browsers and change in rodent taxa in the Pakistani foreland strata¹⁵ and by palaeoceanographic evidence in the western Arabian sea. There, changes in faunal abundance and in rates of organic carbon and opal accumulation indicate the presence of strong upwelling, which today is driven by the northwesterly monsoon winds^{1,9,10}. Together these effects suggest that the monsoon climate system developed in Late Miocene time^{1,6,8,10}.

Where first documented in northern Pakistan, $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ changes in soil carbonates can be dated by comparison with the original, locally determined magnetostratigraphic zonation^{16,17} and calibrated against the most recent magnetic timescale¹⁸. This floral shift was under way by ~7.8–7.6 Myr, and C4 plants dominated the floodplain biomass by 6.0 Myr ago⁶. Similar isotopic shifts in both western and central Nepal^{7,8} have been interpreted as evidence for a coherent and essentially synchronous climate change across the entire Himalayan foothill¹⁴. Two separate magnetostratigraphic studies at Surai Khola^{8,19} in western Nepal indicate that the carbon isotope shift began at ~7.4 Myr ago¹⁸. At Bakiya Khola in central Nepal 250 km farther east, the initiation of the isotope shift occurs at either 6.6 or 5.9 Myr ago^{7,14}, depending on the chosen correlation with the magnetic timescale¹⁸. Thus, although these two sites in Nepal appear to record increases in the importance of C4 plants that are analogous to those observed in Pakistan, the recorded onset of this change becomes younger toward the east; the isotope shift seems to occur 1–1.8 Myr later in central Nepal than in

Pakistan. The apparently diachronous nature of the shift suggests that it should not be considered a precise indicator of monsoonal intensification (which presumably should be a regional, largely synchronous phenomenon). This is consistent with recent interpretations^{10,14} which suggest that the transition from C3 to C4 plants in the Himalayan foreland was not a reflection of regional climate change^{6–8} but rather was part of a worldwide trend in the Late Miocene.

Given that 25% of the modern strontium flux to the world's oceans is contributed by rivers with headwaters in the Himalaya and Tibet²⁰, it has been argued that the observed mid-to-late Cenozoic increase in the Sr isotope content of the world's oceans is largely attributable to chemical weathering in the Himalayan orogen¹². It has been further suggested that this increase was accompanied by or driven by a coeval increase in mechanical weathering^{11,12}, that mean erosion rates have been increasing continuously since 12 Myr ago¹², and that monsoonal intensification in particular should have enhanced overall erosion rates^{5,13}. If so, we might expect to see an isotopic signal of this enhanced Himalayan erosion in the foreland and in the Bengal fan, and we would expect to see increased sediment storage in one or both of these sediment sinks, which are the only reservoirs capable of absorbing the large flux of Himalayan detritus.

Isotopic fingerprinting of potential Himalayan and Tibetan erosional source areas and matching with the isotope characteristics of the Bengal fan strata²¹ clearly show that the High Himalayan crystalline belt was the main contributor to detrital sediments in the fan during the past 17 Myr, and there is no clear evidence for a change in the relative importance of this source area during that time. Detrital mineral ages, interpreted to represent cooling and denudation in rapidly uplifting source areas, show cooling ages remarkably close to depositional ages

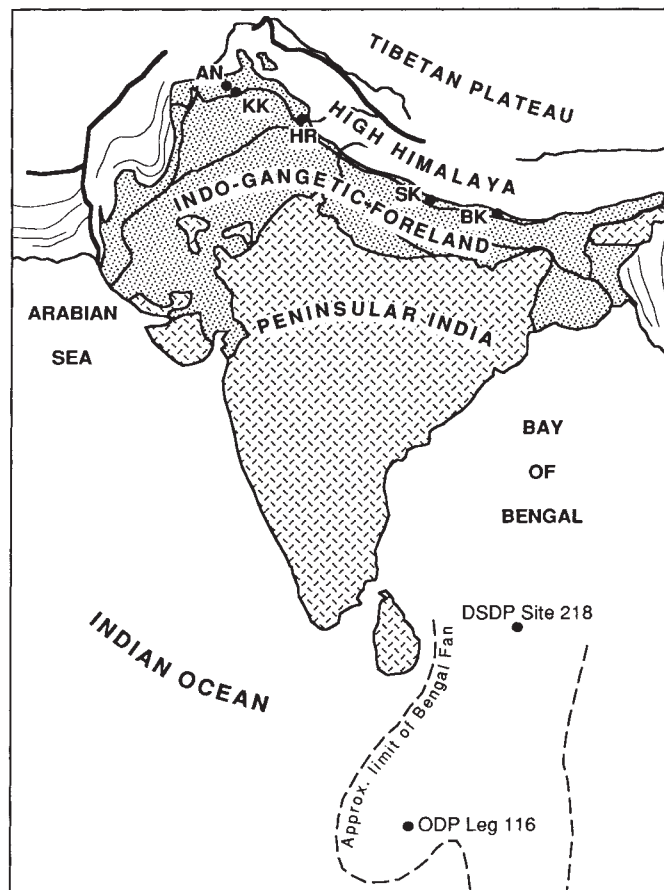


FIG. 1 Location map of Himalayan foreland and Bengal fan showing the locations of marine cores (ODP leg 116, DSDP 218) and foreland magnetostratigraphic sections used in this study. AN: Andar Kas; BA: Bakiya Khola; HR: Haritalyangar; KK: Kotal Kund; SK: Surai Khola.

throughout the Neogene strata of the Bengal fan²², as well as for the Pakistani²³ and Nepali⁷ sectors of the Himalayan foreland. Contrary to predictions of enhanced uplift and erosion rates during the Miocene to present^{11,12}, these isotope data do not indicate significant acceleration in the rates of bedrock uplift and denudation in the High Himalayan source areas in the past 10–15 Myr. It is, however, difficult to generalize about changes in mean erosion rates on the basis of these spatially restricted radiometric data.

A more straightforward approach is to examine the character of deposition and the rate of sediment accumulation in the basins containing the Himalayan detrital record. From ~17–7 Myr ago, the depositional record in the distal Bengal fan (Fig. 1) is marked by rapid sedimentation rates²⁴ and a coarser-grained illite–chlorite-rich assemblage^{21,25}. An interval of accelerated deposition (Fig. 2) is recorded from around 11 Myr ago²⁴. Most of the post 7-Myr interval is dominated by finer-grained smectite–kaolinite-rich strata^{21,25,26}. This compositional change (and possibly the associated grain-size decrease) is interpreted to result not from a change in provenance²¹, but from longer residence on the floodplains of the foreland basin, where the sediment was chemically altered before erosion and redeposition in the Bengal fan.

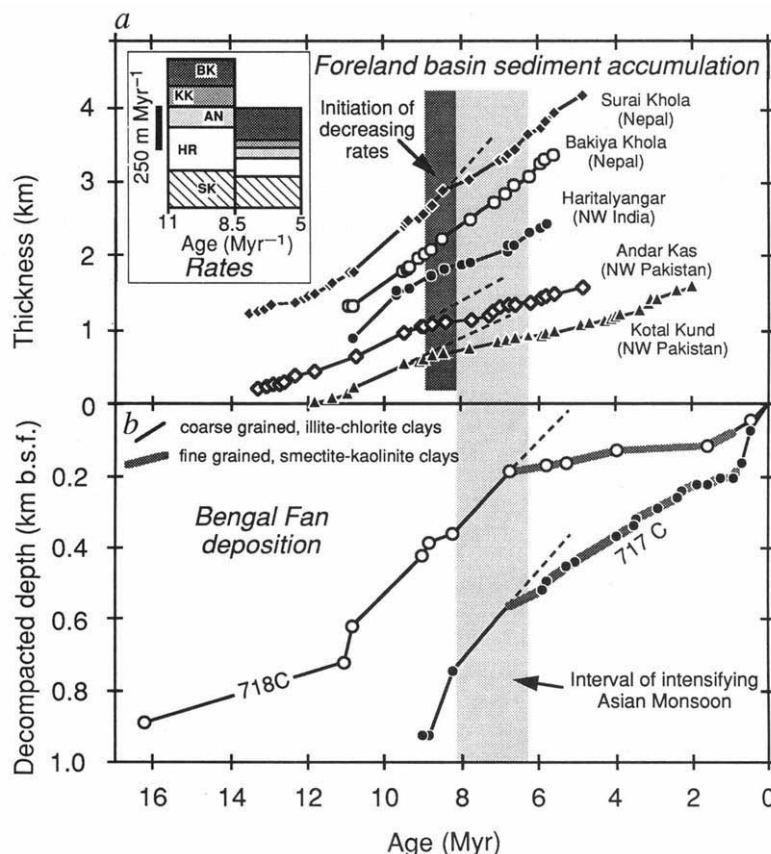
In the Pakistani^{16,17} and Indian²⁷ forelands (Fig. 1), a strong lithostratigraphic contrast is found between strata younger and older than ~8–9 Myr. The older strata (Nagri lithofacies^{16,17} and equivalents) are characterized by thick amalgamated sandstones with reduced preservation of overbank deposits (<30%) that are themselves characterized by deeply leached (1–2 m) palaeosols lacking humic (organic A) horizons⁶. The younger Dhok Pathan¹⁶ lithofacies is characterized by thinner sheet sandstones with abundant overbank deposits (50%) and shallowly leached (<50 cm) palaeosols with increasingly common humic horizons⁶. The contrast between the Nagri and Dhok Pathan lithofacies suggests that fluvial cannibalization of the floodplain was more frequent and that effective annual rainfall (which modulates leaching depths) was considerably greater before 8 Myr ago.

Whereas mean grain-size decreases in the foreland basin, floodplain preservation increases in Dhok Pathan strata in comparison to Nagri strata. These changes are consistent with the interpretations of the finer grain sizes in the Bengal fan after 7 Myr, as well as of concomitant increases in both smectite–kaolinite abundances and the degree of chemical weathering²¹. Although the age (~8 Myr) of the thick-to-thin sandstone transition is very similar in western India²⁷ and northern Pakistan^{16,17}, the coeval interval in Nepal is marked by increasing sandstone thickness, abundance of poorly drained soils, and preservation of organic matter^{14,19}. Consequently, depositional and soil-forming conditions were not coherent across the entire Himalayan foreland during this interval.

The deep-sea cores from the Bengal fan show a significant drop in sediment-accumulation rates (Fig. 2) after ~7 Myr ago²⁴. The succeeding 5–6 Myr of slower deposition correlate with the predominantly smectite–kaolinite assemblage, and they indicate that sediment accumulation in the Bengal fan decreased²¹ at a time when mechanical erosion of the Himalaya is suggested to have increased^{5,11,13,28}. One possible solution to this apparent contradiction would be to increase the amount of sediment storage within the Himalayan foreland. But sediment-accumulation rates synthesized from across the foreland basin based on magnetostratigraphic data^{7,16,17,19,27,29–31} rule this out (Fig. 2). Although sediment accumulation accelerated in many localities between 11.5 and 10.5 Myr ago (also in the Bengal fan), nearly all sections show either steady or diminishing rates of accumulation for strata <8.5 Myr old.

It could be argued that the accumulation-rate changes on the Bengal fan that are presented here (Fig. 2) are not representative of the entire fan, or are the result of sea-level changes rather than sediment delivery. Although less well constrained, the only other relevant deep-sea cores (DSDP site 218, Fig. 2) show age–depth trends consistent with coeval decreases in accumulation rates across the fan. Neither the timing and magnitude of sea-level variations nor processes of lobe switching can adequately account for the observed mineralogical and rate changes²¹.

FIG. 2 a, Sediment-accumulation curves from the Himalayan foreland basin based on magnetostratigraphic time control^{7,8,16,19,27}. Overall, accumulation rates (equal to the slope of the accumulation curve and shown in the inset) decrease after 8.5 Myr ago, and they remain steady or lower during the subsequent interval of intensifying Asian monsoon. Dashed lines show constant accumulation for comparison. AN: Andar Kas; BA: Bakiya Khola; HR: Haritalyangar; KK: Kotal Kund; SK: Surai Khola. b, Decompressed sediment-accumulation rates for the cores of ODP leg 116 (sites 717, 718) in the Bengal fan (b.s.f., below sea floor). Stratigraphic and palaeontological control from refs 21, 24, 25. Leg 116 sites show decreasing sedimentation rates beginning ~6.8 Myr ago (dashed lines show constant rates for comparison). The change in sedimentation rate corresponds with a major change in lithology, mean grain size and clay mineralogy which can be correlated at sites 717, 718 and also DSDP 218 (refs 21, 25, 36–38). This change is best dated at site 718, and is inferred to have occurred essentially synchronously at sites 717 and 218 (ref. 21). Rates remain low until the coarse grained illite–chlorite facies returns ~0.8 Myr ago.



It could also be argued that apparent rate changes in the foreland are strongly dependent on either the choice of magnetic timescales or on the precision of measurements of stratigraphic sections. Although we use the recent magnetic timescale of Cande and Kent¹⁸ here, comparable changes in rates with slight shifts in the absolute age of inflections (<0.6 Myr) are obtained based on other widely used timescales^{32,33}. Similarly, to reverse the observed slowing of sediment-accumulation rates, either the pre- or post-9-Myr strata would have had to have been systematically mismeasured by >20%, which is considerably greater than the uncertainties associated with these measured sections. Therefore, the calculated rate changes form a firm basis for evaluating sediment storage within the foreland.

Consequently, a consistent picture can be developed for sediment storage in both the Himalayan foreland basin and the Bengal fan for middle and late Miocene deposition. Faster rates of sediment accumulation, coarser mean grain size and shorter residence times on the foreland floodplain in middle Miocene times are succeeded both by finer mean grain size and reduced volumes of sediment storage in both the foreland and the Bengal fan and by increased floodplain residence times in the Late Miocene beginning 8–7 Myr ago. This change to a regime of lower detrital sediment fluxes to the foreland and deep-sea basins corresponds with the timing of apparent intensification of the Asian monsoon and with the increased importance of chemical relative to mechanical denudation²¹. The terrestrial evidence (soil-carbonate isotopes) for this climate change seems moderately diachronous across the foreland. It might be expected that monsoonal intensification would result in increased sediment production and delivery. However, there is no evidence for a concurrent increase in mechanical weathering, detrital sediment delivery, or storage of detrital sediments in the adjacent sedimentary basins. It is possible that, despite increased mechanical erosion, more sediment was stored within the Himalaya itself, rather than being discharged to the foreland. But there is no evidence to support this hypothesis, and in any event, it is doubtful that the Himalaya could have stored enough sediment to account for the long-term changes seen in sediment fluxes downstream.

The cause of the reduced sediment flux to the late Miocene basins is not known. At least three alternative hypotheses are possible. Rates of tectonically induced bedrock uplift may have declined, contributing to overall lower relief and detrital sediment delivery. The development of the monsoon climate may have been associated with decreased glaciation in the Himalaya^{1,34}, leading to decreased mechanical erosion, but increased chemical weathering rates. Vegetation changes driven by the enhanced monsoon may have caused faster chemical weathering due to changing CO₂ levels in soils, whereas the same vegetation may have stabilized slopes and decreased the detrital sediment flux. If slope stabilization did result from dense plant cover at higher elevations, this would imply that the transition from C3 to C4 plants was confined to lower altitudes. Isotopic evidence of transition from C4 to C3 vegetation with increasing elevation has been observed elsewhere³⁵. It is perhaps worth noting that today much of the Lesser Himalaya and Siwalik hills in India and Nepal are covered in jungle, despite the strongly seasonal climate. That the development of a monsoonal climate should be associated with decreased sediment yields is perhaps surprising. However, given the present state of knowledge about the functional relationships governing sediment production, there may be other surprises in store for students of the interplay between tectonism, climate and erosion. □

5. Rea, D. K. *Eos* **72**, 257 (1991).
6. Quade, J., Cerling, T. E. & Bowman, J. R. *Nature* **342**, 163–166 (1989).
7. Harrison, T. M. *et al.* *J. Geol.* **101**, 157–175 (1993).
8. Cater, J. M. L., *et al.* in *Himalayan Tectonics* (eds. Treloar, P. J. & Searle, M.). (Geol. Soc. London spec. Publ. (in the press)).
9. Kroon, D., Steens, T. & Troelestra, S. R. in *Proc. ODP scient. Res.* **117**, 257–264 (1991).
10. Prell, W. L., Murray, D. W., Clemens, S. C. & Anderson, D. M. in *The Indian Ocean: A Synthesis of Results from the Ocean Drilling Program* (Am. geophys. Un. in the press).
11. Raymo, M. E. & Ruddiman, W. F. *Nature* **359**, 117–122 (1992).
12. Richter, F. M., Rowley, D. B. & DePaolo, D. J. *Earth planet. Sci. Lett.* **109**, 11–23 (1992).
13. Raymo, M. E., Green, P. F. & George, A. *Eos* **72**, 257 (1991).
14. Quade, J., Cater, J. M., Ojha, T. P. & Harrison, T. M. *Geol. Soc. Am. Bull.* (submitted).
15. Barry, J. C., Johnson, N. M., Raza, S. M. & Jacobs, L. L. *Geology* **13**, 637–640 (1985).
16. Johnson, N. M., Opdyke, N. D., Johnson, G. D., Lindsay, E. H. & Tahirkehi, R. A. K. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **37**, 17–42 (1982).
17. Johnson, N. M., Stix, J., Tauxe, L., Cervený, P. F. & Tahirkehi, R. A. K., *J. Geol.* **93**, 27–40 (1985).
18. Cande, S. C. & Kent, D. V. *J. geophys. Res.* **97**, 13917–13953 (1992).
19. Appel, E., Rosler, W. & Corvinus, G. *Geophys. J. int.* **105**, 191–198 (1991).
20. Palmer, M. R. & Edmond, J. M. *Earth planet. Sci. Lett.* **92**, 11–26 (1989).
21. France-Lanord, C., Derry, L. & Michard, A. in *Himalayan Tectonics* (eds Treloar, P. J., and Searle, M.). (Geol. Soc. London spec. Publ. in the press).
22. Copeland, P. & Harrison, M. T. *Geology* **18**, 354–359 (1990).
23. Cervený, P. F., Naeser, N. D., Zettler, P. K., Naeser, C. W. & Johnson, N. M. in *New Perspectives in Basin Analysis* (eds Kleinspehn, K. L., Paola, C. & Bouma, A.) (Springer, New York, 1988).
24. Gartner, S. in *Proc. ODP scient. Res.* **116**, 165–187 (1990).
25. Shipboard Scientific Party, Sites 717, 718 in *Proc. ODP init. Rep.* **116** (1989).
26. Bouquillon A., France-Lanord C., Michard A & Tiercelin J.-J. (1990) in *Proc. ODP scient. Res.* **116**, 43–58 (1990).
27. Johnson, G. D., Opdyke, N. D., Tandon, S. K. & Nanda, A. C. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **72**, 223–249 (1983).
28. Raymo, M. E., Ruddiman, W. F. & Froelich, P. N. *Geology* **16**, 649–653 (1988).
29. Reynolds, R. G. H. thesis, Dartmouth College (1980).
30. Reynolds, R. G. H. & Johnson, G. D. in *The Chronology of the Geologic Record* (ed. Snelling, N. J.) (Geol. Soc. London Mem. **10**, 1985).
31. Tauxe, L. & Opdyke, N. D. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **37**, 43–61 (1982).
32. Berggren, W. A., Kent, D. V., Flynn, J. J. & van Couvering, J. A. *Geol. Soc. Am. Bull.* **96**, 1407–1418 (1985).
33. Harland, W. B., *et al.* *A Geologic Time Scale* (Cambridge Univ. Press, 1990).
34. Le Fort, P. in *Tectonic Evolution of the Tethyan Regions* (ed. Sengor, A.M.C.) (Proc. NATO ASI, Reidel, 1985).
35. Livingstone D. A. and Clayton W. D. *Quat. Res.* **13**, 392–402 (1980).
36. Shipboard Scientific Party, Site 218 in *Init. Rep. DSDP* **12**, (1974).
37. Bouquillon A., Chamley, H. & Frolich, F. *Oceano Acta* **12**, 133–147 (1989).
38. Pimm, A. C. in *Init. Rep. DSDP* **22**, (1974).

ACKNOWLEDGEMENTS. The contributions of D.W.B. to this research were funded in part by National Geographic Research and the NSF. Those of L.A.D. and C.F.L. were carried out under the auspices of the French programme "Dynamique et Bilans de la Terre fluviale et érosion."

Timing of Tibetan uplift constrained by analysis of volcanic rocks

Simon Turner, Chris Hawkesworth, Jiaqi Liu*, Nick Rogers, Simon Kelley & Peter van Calsteren

Department of Earth Sciences, The Open University, Walton Hall, Milton Keynes MK7 6AA

*UK Institute of Geology, Chinese Academy of Sciences, PO Box 634, Beijing 100011, China

THE Tibetan plateau has had a central role in the development of recent models for the mechanics of mountain belts^{1,2} and Cenozoic global climate change³. The present elevation and extensional deformation of the plateau probably result from uplift owing to convective thinning of the underlying lithospheric mantle^{1,2}. An age for the uplift would provide a valuable constraint on these models; but because recently proposed indicators of uplift are all climate-dependent, they are equivocal, possibly reflecting global cooling rather than regional uplift⁴. Here we present new geochemical data on post-mid-Miocene lavas from the plateau, which show that the lavas were derived from the lithospheric mantle. Simple thermal arguments indicate that the generation of such magmas also necessitates thinning of the lithospheric mantle. Thus volcanism is coincident with uplift, providing a climate-independent means of dating the onset of uplift. Dating by the laser ⁴⁰Ar/³⁹Ar technique places the beginning of this volcanism, and therefore the time of uplift of the Tibetan plateau, at 13 Myr ago.

Received 16 February; accepted May 1993.

1. Prell, W. L. & Kutzbach, J. E. *Nature* **360**, 647–652 (1992).
2. Harrison, T. M., Copeland, P., Kidd, W. S. F. & Yin, A. *Science* **255**, 1663–1670 (1992).
3. Ruddiman, W. F. & Raymo, M. E. *Phil. Trans. R. Soc.* **B318**, 411–430 (1988).
4. Ruddiman, W. F. & Kutzbach, J. E. *J. geophys. Res.* **94**, 18409–18427 (1989).

The Tibetan plateau (Fig.1) formed in response to the collision of India and Asia around 50–45 Myr ago. Since then, India has moved northwards with respect to Asia at a constant rate ($4\text{--}5\text{ cm yr}^{-1}$), apparently accommodated by internal deformation (see for example ref. 5) of India and Asia, and seismic data indicate that the crust has reached twice normal thickness ($\sim 70\text{ km}$)⁶. The plateau has an average elevation of $5\text{--}5.5\text{ km}$ and is currently extending east–west ($\sim 1\%$ Myr^{-1}) on normal faults^{7–10}. In a convergent orogen the build-up of topography results in increased gravitational potential energy and extensional buoyancy forces^{1,2,11–13} but clearly plate motions can only build up topography to the point where the resultant buoyancy forces balance these driving forces. It follows that extension only occurs if the plate forces are reduced, or if the topography can be raised by some other means such that the buoyancy forces exceed the external driving forces. Therefore it has repeatedly been argued that, as there is no evidence for a reduction in the rate of convergence between India and Asia, the current extension on the plateau is most plausibly explained by an increase in elevation resulting from removal of part of the underlying lithospheric mantle^{1,2}.

The notion that the behaviour of the crust and mantle lithosphere may become decoupled during deformation has considerable implications for the evolution of mountain belts^{1,2,11–13}. In addition, high plateaus affect global atmospheric circulation, and thus it has been suggested that uplift of the

Tibetan plateau may be the cause of Cenozoic global cooling and increases in Atlantic $\delta^{18}\text{O}$ values³. Consequently, much effort has gone into trying to date the uplift of the plateau, using either palaeobotanical data¹⁴ to record altitude changes or sedimentological and mineral cooling age data to record rapid increases in exhumation^{15–18}. Unfortunately, exhumation does not necessarily equate to surface uplift¹⁹ and, as data on exhumation come from the plateau margins, they provide little information on the plateau itself (which does not display exhumation). Moreover, as erosion and sedimentation rates are related to climate and the correlation of fossil flora with altitude depends on palaeotemperature, these data might equally reflect global cooling⁴. In summary, there is currently little reliable information on the age of uplift of the plateau. As an alternative approach, we have looked at its punctuated magmatic history.

Changes in tectonic regime are typically marked by changes in the composition of the associated magmatism. Before collision, subduction of oceanic crust was accompanied by calc-alkaline arc magmatism in the High Himalaya^{20,21}. Small volumes of crustally derived leucogranites were subsequently emplaced here at $\sim 20\text{ Myr}$ ago, possibly resulting from decompression associated with the onset of rapid exhumation²². In contrast, the Tibetan plateau was generally devoid of magmatic activity between 90 Myr ago and the mid-Miocene, since which time it has become the site of widespread (Fig.1), but small-volume, eruptions of potassic basalts and lesser felsic magmas^{21,23–25}.

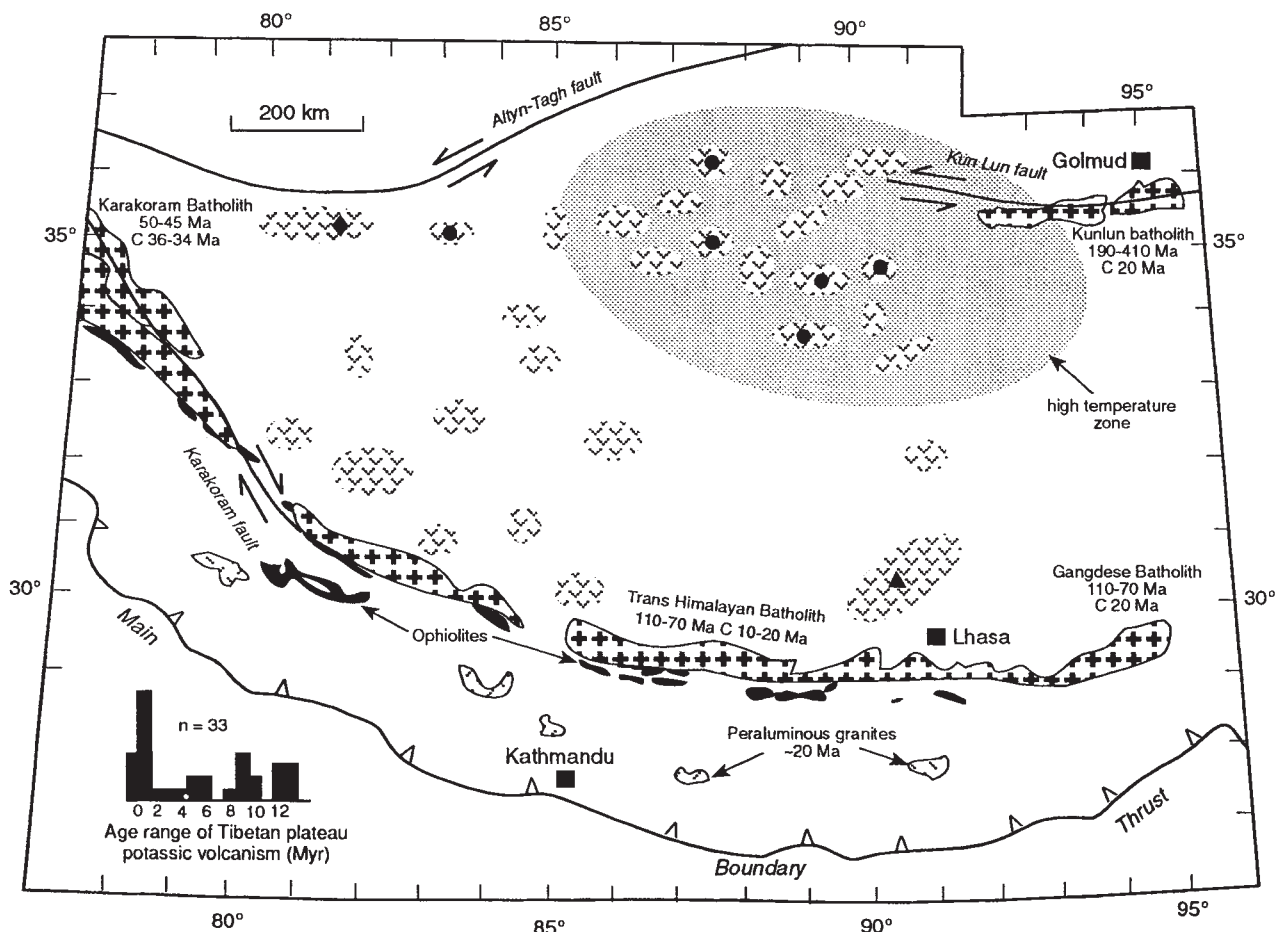


FIG. 1 Simplified map (based on refs. 8, 23) of Tibet and the Himalaya showing subduction-related calc-alkaline intrusives before and during collision and the suture (defined by ophiolites) marking the southern border of the Tibetan plateau. Intrusive and cooling ages (marked "C") are given for some of these batholiths. Post-mid-Miocene volcanics on

the Tibetan Plateau are shown schematically with the inset (bottom left) giving the currently available $^{40}\text{Ar}/^{39}\text{Ar}$ age determinations. Sampling localities: circles, this study or ref. 24; triangles, ref. 23; diamonds ref. 25. The high-temperature zone⁶ is also shown.

TABLE 1 Selected analyses of post Mid-Miocene Tibetan volcanics

Sample number ⁴⁰ Ar/ ³⁹ Ar age	K9007 13.3	K9024 13.0	K9031 11.8	K9038 9.6	K9008 9.4	K9021 8.5	K89G185 5.6	K89G200 3.6	K738 0.5	K732 0.3
SiO ₂	59.44	56.57	55.95	55.31	61.34	62.02	51.42	56.67	55.44	56.22
TiO ₂	1.54	1.58	1.55	1.59	1.62	1.57	1.85	1.43	2.32	2.28
Al ₂ O ₃	15.30	15.41	15.19	15.63	15.15	15.14	14.62	13.80	14.12	14.54
Fe ₂ O ₃ *	6.13	7.05	6.86	5.88	6.22	5.92	9.61	7.31	11.04	8.78
MnO	0.09	0.10	0.10	0.15	0.09	0.08	0.14	0.10	0.14	0.11
MgO	1.87	3.21	3.35	2.53	1.23	1.77	4.03	4.37	3.67	2.90
CaO	4.75	5.64	6.30	7.38	3.77	3.99	7.59	6.35	6.33	5.14
Na ₂ O	3.42	3.69	3.61	3.63	3.82	3.77	4.60	4.02	2.56	2.84
K ₂ O	3.93	3.81	3.83	3.71	4.21	4.12	4.77	5.40	3.39	4.04
P ₂ O ₅	0.78	0.91	0.82	0.83	0.84	0.84	-	-	1.12	1.23
Loss on ignition	2.97	1.82	1.99	3.26	0.94	0.81	2.09	0.32	1.10	1.46
Total	100.22	99.79	99.55	99.90	99.23	100.03	100.72	99.77	101.23	99.54
Cr	15	86	62	109	21	15	88	83	73	239
Ni	21	51	39	55	21	21	72	61	45	51
Sc	8	15	14	15	7	10	10	9	14	12
V	92	117	116	133	96	96	-	-	133	121
Pb	67	32	30	21	29	39	29	35	42	34
Rb	198	118	119	98	214	211	122	198	108.7	137.4
Sr	686	1,025	983	1,075	713	686	2,011	1,567	913.9	1,101
Ba	1,387	1,635	1,626	1,315	1,466	1,389	2,563	2,076	2,206	1,985
Ga	22	19	18	20	22	23	22	20	20	16
Nb	39	39	39	36	39	39	31	26	44	38
Zr	504	433	433	328	507	514	411	407	55	461
Y	28	26	25	25	26	27	32	23	32	28
Th	31	23	22	9	33	32	27	44	33	27
U	5	3	3	0	5	6	2.8	6.8	3	2
La	136	122	116	82.4	133	137	177	161	148	170
Ce	255	233	224	163	255	258	333	287	260	326
Nd	97.9	90.3	86.5	67.7	97.5	95.6	142	119	129.2	154.2
Sm	13.6	12.3	11.8	9.75	13.4	13.1	20.2	16.1	19.2	21.6
Eu	2.91	2.92	2.82	2.56	2.9	2.86	4.3	3.6	2.6	3.2
Tb	1.10	1.00	0.96	0.93	1.10	10.5	1.30	1.00	0.84	1.20
Yb	2.13	1.9	1.75	1.88	1.95	2.11	1.5	1.4	1.2	1.3
⁸⁷ Sr/ ⁸⁶ Sr	0.70918	0.70837	0.70916	0.70803	0.70910	0.70916	0.70808	0.70819	0.71005	0.71054
⁸⁷ Sr/ ⁸⁶ Sr (t)	0.70906	0.70831	0.70910	0.70799	0.70899	0.70904	0.70807	0.70817	0.71005	0.71054
¹⁴³ Nd/ ¹⁴⁴ Nd	0.51229	0.51232	0.51233	0.51229	0.51227	0.51231	0.51233	0.51231	0.51237	0.51222
ε _{Nd} (t)	-6.6	-5.9	-5.8	-6.6	-7.2	-6.3	-5.9	-6.3	-5.3	-8.1
T _{DM}	0.9	0.9	0.9	1.0	1.0	0.9	0.9	0.9	0.9	1.0

All analyses done at the Open University (K89G185, 200 from ref.25). NBS987 ratio of ⁸⁷Sr/⁸⁶Sr=0.710252±16 (*n*=5); J&M ratio of ¹⁴³Nd/¹⁴⁴Nd=0.511794±9 (*n*=13). Argon age determinations made by laser Ar-probe with methods and data reduction as detailed in ref.30. ⁸⁷Sr/⁸⁶Sr (t) and ε_{Nd}(t) refer to values calculated at the argon age of the sample.

Table 1 summarises our database, and the accompanying laser ⁴⁰Ar/³⁹Ar age determinations which show that these lavas have been erupting since 13 Myr ago.

The dominant lavas are clinopyroxene and plagioclase-phyric, potassic basaltic andesites (shoshonites) which are extremely enriched in the incompatible elements (Fig. 2a) with Sr 800–2,500 p.p.m., Ba 1,000–2,000 p.p.m., La > 100 p.p.m., and K₂O concentrations around 3–4.5% (Table 1). The high Ti/Y and low Rb/Ba ratios are features not generally ascribed to the continental crust (Fig. 2b, inset) showing that the compositions do not arise from crustal contamination of asthenospheric magmas. The incompatible behaviour of Sr and the absence of significant Eu anomalies indicate that the source was plagioclase-free and therefore probably located in the mantle, yet the low Nb/La ratios, high radiogenic Sr (~0.7105) and unradiogenic Nd (ε_{Nd} ~ -7) show that the source was distinct from the convecting asthenosphere (Fig. 2b). Similarly, the degrees of LREE enrichment and REE fractionation are too great for the magmas to have been generated from a depleted

mantle source by a single melting episode. The likely explanation is a two-stage model in which small-degree partial melts infiltrate and enrich the subcontinental lithospheric mantle, which is then remobilized at a later stage. Because small partial melts have elevated H₂O, and high Rb/Sr and Nd/Sm ratios, this enriched lithospheric mantle will be hydrous with a lowered solidus and, over time, will develop lower ε_{Nd} and higher ⁸⁷Sr/⁸⁶Sr than the underlying asthenosphere. The isotopic data require that incompatible-element enrichment of this source was long-lived, with Nd model ages (T_{DM}) being ~ 900 Myr (Table 1).

The appearance of widespread, lithospheric mantle-derived volcanism on the plateau marks a change in the tectonic regime and requires a thermal explanation. Melting cannot reflect heating of thickened crust, because the low conductivity of rocks means increased radiogenic heating due to thickening will develop only on the timescale of the thermal time constant of the thickened lithosphere (~120 Myr). The thickening of the lithosphere and low degrees of extension also rule out the

possibility that the magmas reflect decompression melting of the asthenosphere²⁶; and the areal extent and amount of uplift are too great, and the magma volumes too small, to be attributed to the effects of a mantle plume. The extension of the volcanism, 1,000 km north of the arc batholiths rules out a subduction origin; furthermore, magmatism occurred 30 Myr after collision, a delay that would not be expected if oceanic or continental subduction²⁵ were the trigger. This timing does compare well with the lag predicted for lithospheric thinning, which will proceed rapidly once initiated²⁷. Lithospheric thinning has been discussed previously²⁵ but was considered to be of little importance relative to continental subduction and dehydration. In addition to the timing problems mentioned above, fluid release accompanying the physically improbable subduction of buoyant continental lithosphere provides no thermal rationale for melting within an already hydrous, lithospheric mantle. We conclude that, in Tibet, the only plausible means of attaining temperatures high enough for melting within the lithospheric mantle is by thinning it, a mechanism that is consistent with recent seismic observations⁶.

Our preferred model is illustrated in Fig. 3. The middle part of

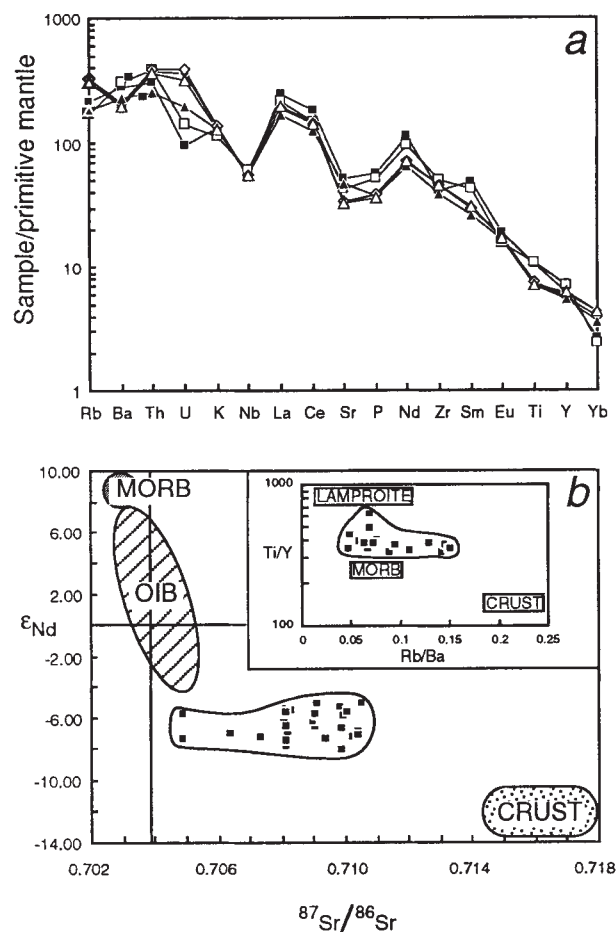


FIG. 2 *a*, 'Spidergram' of selected post-mid-Miocene basaltic andesites from the Tibetan plateau, normalized to primitive mantle. The patterns show extreme incompatible-element enrichment not typical of crustal rocks or asthenospheric mantle melts such as mid-ocean-ridge or ocean-island basalts. *b*, Initial Nd and Sr isotope ratios which indicate that these small degree melts were isolated from mantle convection for long enough to develop distinct isotopic anomalies. The inset of Ti/Y against Rb/Ba shows that the elemental composition of the lavas is more akin to small-degree mantle melts in equilibrium with residual garnet (represented by lamproites) than to crustally contaminated mantle magmas.

the lithospheric mantle contains a hydrous, K-rich layer, which is the zone where small partial melts freeze²⁸ (Fig. 3*b*). Before collision, the plateau had zero elevation as indicated by Eocene fossiliferous limestones²⁹. Following collision, homogeneous thickening by a factor of two (to obtain 70-km-thick crust) resulted in an isostatically supported elevation of 3–4 km. The

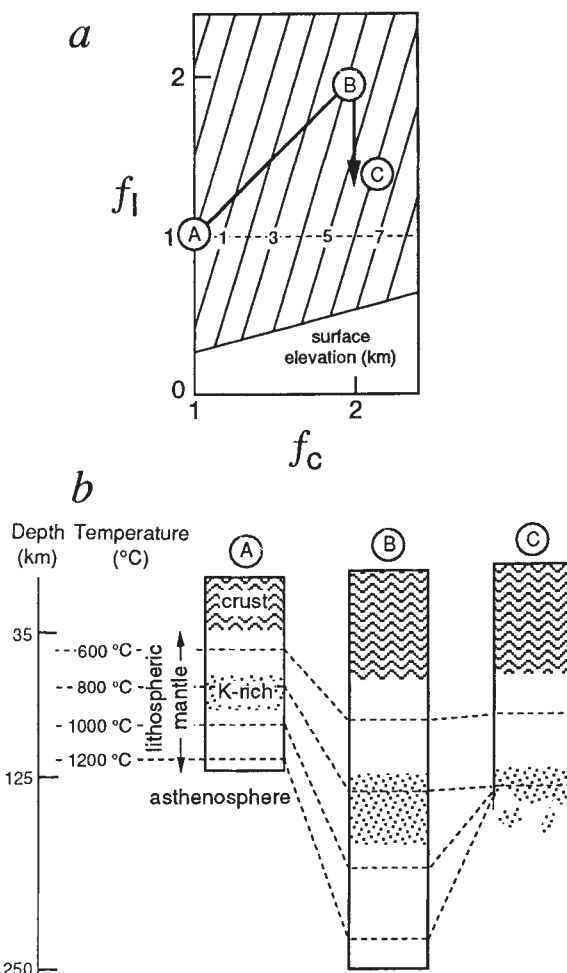


FIG. 3 Thermal and isostatic modelling of a three-stage evolution for the Tibetan plateau, with collisional thickening from *a* to *b* and thinning of the lithospheric mantle with attendant magmatism and extension from *b* to *c*. *a*, f_c-f_1 plane (where f_c is the crustal thickening factor and f_1 is the total lithospheric thickening factor (see ref. 11) (parameterization and input conditions from refs 11, 13) contoured for the isostatically compensated surface elevation. Point A represents crust of normal thickness at sea level before collision. Homogeneous thickening ($f_c=f_1$) occurs until B, when the whole lithosphere reaches twice its original thickness and has an elevation of 3–4 km. From B to C the mantle lithosphere is thinned (~60%) to obtain its current elevation of 5–5.5 km. This increase in elevation correspondingly increases the gravitational potential energy of the plateau, promoting extensional deformation. *b*, A depth-temperature profile for the three stages, illustrating the accompanying thermal history. Initially the lithosphere is of normal thickness and geotherm and contains a volatile-bearing 'K-rich' layer with a solidus between 750 and 900 °C. Homogeneous thickening (B) moves all material points downwards. Subsequent thinning of the lithospheric mantle compresses the geotherms in the lower portion of the lithospheric mantle and brings the lower half of the K-rich layer into contact with asthenosphere, passing through its solidus and producing potassic basaltic andesites. Conductive heating will result in relaxation of the geotherm on a timescale dictated by the thermal time constant of the lithosphere (~120 Myr) such that the entire K-rich layer will pass through its solidus over ~10 Myr.

lithospheric mantle was then thinned by ~60% to achieve the observed 5–5.5 km elevation, the increase in gravitational potential energy produced by this topography being of the order required to induce extension^{1,2,11–13}. An important point is that the thermal profile after thinning (Fig. 3b) puts 50% of the lithospheric mantle above the hydrous peridotite solidus, so the K-rich layer, insulated when the lithosphere was thick, began to melt. About 10 million years is required (see Fig. 3b legend) for solidus temperatures to be exceeded throughout the K-rich layer, in accordance with the observed long period of magmatism. One outcome of the geometry in Fig. 3a is that areas of thinner crust require greater degrees of thinning to maintain an elevation of 5.5 km. The area of thinnest crust on the plateau is the zone of inferred high mantle temperatures⁶ (Fig. 1) and contains the greatest concentration of volcanic rocks.

Thus uplift, extension and lithospheric-mantle-derived magmatism may all be reconciled with thinning of the lithospheric mantle¹³. Assuming that the samples dated represent the age range of the province, our present data place uplift of the Tibetan plateau at 13 Myr. Because extensional deformation is also a response to the increased elevation of the plateau, the timing of the onset of extension provides a rough check of this uplift age. Thrust faults indicate that the plateau was under compression during most of the Tertiary, and recent dating suggests that east–west extension commenced around 13–10 Myr ago^{8–10}, in excellent agreement with the age determined from magmatism. □

Received 10 December 1992; accepted 18 May 1993.

1. England, P. C. & Houseman, G. A. *Phil. Trans. R. Soc.* **326**, 301–320 (1988).
2. England, P. C. & Houseman, G. A. *J. geophys. Res.* **94**, 17561–17579 (1989).

3. Raymo, M. E. & Ruddiman, W. F. *Nature* **359**, 117–122 (1992).
4. Molnar, P. & England, P. *Nature* **346**, 29–34 (1990).
5. Kidd, W. S. F. & Molnar, P. *Phil. Trans. R. Soc.* **327**, 337–363 (1988).
6. Molnar, P. *Earth planet. Sci. Lett.* **101**, 68–77 (1990).
7. Armijo, R., Tapponnier, P., Mercier, J. L. & Han Tong Lin *J. geophys. Res.* **91**, 13803–13872 (1986).
8. Pan, Y. & Kidd, W. S. F. *Geology* **20**, 775–778 (1992).
9. Mercier, J.-L., Armijo, R., Tapponnier, P., Carey-Gailhardis, E. & Han Tong Lin *Tectonics* **6**, 275–304 (1987).
10. Maluski, H., Matte, P. & Brunel, M. *Tectonics* **7**, 299–326 (1988).
11. Sandiford, M. & Powell, R. *Earth planet. Sci. Lett.* **98**, 154–165 (1990).
12. Zhou, S. & Sandiford, M. *J. geophys. Res.* **97**, 14207–14223 (1992).
13. Turner, S. P., Sandiford, M. & Foden, J. D. *Geology* **20**, 931–934 (1992).
14. Xu R. *Geological and Ecological studies of Qinghai-Xizang Plateau* **1**, 139–144 (Beijing Science 1981).
15. Copeland, P., Harrison, T. M., Kidd, W. S. F., Xu, R. & Zhang, Y. *Earth planet. Sci. Lett.* **86**, 240–252 (1987).
16. Richter, F. M., Lovera, O. M., Harrison, T. M. & Copeland, P. *Earth planet. Sci. Lett.* **105**, 266–278 (1991).
17. Burbank, D. W. *Nature* **357**, 680–683 (1992).
18. Amano, K. & Taira, A. *Geology* **20**, 391–394 (1992).
19. England, P. C. & Molnar, P. *Geology* **18**, 1173–1177 (1990).
20. Harris, N. B. W., Xu, R., Lewis, C. L. & Jin C. *Phil. Trans. R. Soc.* **327**, 145–168 (1988).
21. Pearce, J. A. & Mei H. *Phil. Trans. R. Soc.* **327**, 169–201 (1988).
22. Harris, N. B. W., Inger, S. & Massey, J. *J. geol. Soc. Land. Spec. Publ.* (in the press).
23. Coulon, C., Maluski, H., Bollinger, C. & Wang, S. *Earth planet. Sci. Lett.* **79**, 281–302 (1986).
24. Liu, J. & Maimaiti, Y. *Bull. Glac. Res.* **7**, 187–190 (1989).
25. Arnaud, N. O., Vidal, Ph., Tapponnier, P., Matte, Ph. & Deng, W. M. *Earth planet. Sci. Lett.* **111**, 351–367 (1992).
26. McKenzie, D. & Bickle, M. J. *J. Petrol.* **29**, 625–679 (1988).
27. Houseman, G. A., McKenzie, D. P. & Molnar, P. *J. J. geophys. Res.* **86**, 6115–6132 (1981).
28. McKenzie, D. P. *Earth planet. Sci. Lett.* **95**, 53–72 (1989).
29. Norin, E. *Geological Explorations in Western Tibet*. (Pub 29, Tryckeri Aktiebolaget, Thule, 1947).
30. Hawkesworth, C. J. et al. *J. geol. Soc. Lond. spec. Vol.* **26**, 221–241.

ACKNOWLEDGEMENTS. We thank N. Harris, M. Sandiford, P. England, N. Arnaud and S. Reddy for comments and discussions. Funding for isotope research at the Open University is partly provided by NERC. J. L. thanks the Chinese Academy of Sciences and the British Royal Society for financial assistance. Reviews by Julian Pearce and Phil England helped us to strengthen our arguments.

Pressure-induced coordination changes of transition-metal ions in silicate melts

Hans Keppler & David C. Rubie

Bayerisches Geoinstitut, Universität Bayreuth,
W-8580 Bayreuth, Germany

THE strong partitioning of cobalt and nickel from silicate melts into solid silicate and metal phases^{1–3} during the large-scale fractionation processes that took place early in Earth history^{4–9} determines their abundance in mantle-derived magmas and in the metallic core. The global distributions of these elements have therefore been used to constrain models of the Earth's chemical evolution^{4–8}, as well as that of the Moon^{10,11}. But in virtually all model calculations^{4,6,7,10,11} the effect of pressure on partitioning has been neglected — these models have used partition coefficients determined at very low pressures. Here we present crystal-field spectra of doped silicate glasses quenched from high-pressure, high-temperature melts, which show that the coordination of cobalt and nickel ions in silicate melts changes at high pressures. Our results suggest that the partitioning of these species into the melt phase increases markedly with pressure, such that the crystal/melt and metal/melt partition coefficients decrease by more than an order of magnitude at lower-mantle conditions.

The speciation of divalent cobalt and nickel in quenched silicate melts (glasses) can easily be studied by crystal field spectroscopy^{12–14} which measures the energy of electronic transitions between the *d*-orbitals of these ions. At atmospheric pressure, Co²⁺ seems to be in tetrahedral coordination in silicate glasses^{13,14}. The environment of Ni²⁺ is very distorted from an

ideal octahedral geometry¹³. A distinct five-coordinated site has been suggested¹⁵, but a distorted octahedral geometry can equally well account for the main spectroscopic features of Ni-doped glasses. In addition, some tetrahedral Ni²⁺ might be present in some glass compositions¹², whereas in other glasses the amount of tetrahedral Ni²⁺ seems to be small or negligible¹³. Spectroscopic studies of nickel and cobalt in glasses can also help to explain the partition behaviour of these elements. The partitioning of Co and Ni between minerals and silicate melts is essentially determined by the different crystal field stabilization energies (CFSE) of Co and Ni in crystals and melts¹³. Quantitative predictions of partition coefficients can therefore be based on spectroscopic measurements of CFSE. The fact that the CFSE of Co and Ni in melts is much smaller than in most minerals is directly related to the different coordination geometries in these phases. This means that any pressure-induced coordination change will strongly affect the partition coefficients of these elements.

To investigate the speciation of Ni and Co in silicate melts at high pressure, we measured the crystal field spectra of albite (NaAlSi₃O₈) glasses doped with 1 wt% NiO or CoO and quenched from melts at up to 2,900 °C and 100 kbar. We used albite because it is easily quenched to a glass. Studies at 1 atm show that the speciation of Ni and Co in albite glass is a good model for a wide range of silicate melt compositions, including those similar to basalt^{12–14}. Experiments were carried out in a multi-anvil apparatus using MgO as the pressure-transmitting medium, rhenium capsules, LaCrO₃ heaters and W/Re thermocouples. After reaching the desired pressures, the charges were heated to the desired temperature and quenched immediately (to room temperature) at constant pressure. To avoid oxidation of the Re capsules, a trace of Co or Ni metal was added to the charges. The run products were completely glassy and free of crystals. Electron microprobe analyses showed that their composition was unchanged and, in particular, there was no

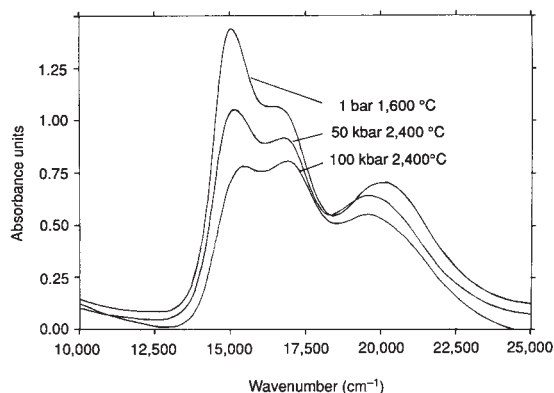


FIG. 1 Crystal field spectra of Co-doped albite glasses quenched from different pressures and temperatures (1 wt% CoO, renormalized to 0.5 mm thickness). Spectra were obtained from polished glass discs using a Bruker IFS 120 Fourier transform infrared spectrometer (Xenon arc lamp or tungsten light source, quartz beamsplitter, Si-detector). The Co-distribution in the sample quenched from 50 kbar is slightly inhomogeneous and extinction coefficients should not be calculated from this spectrum.

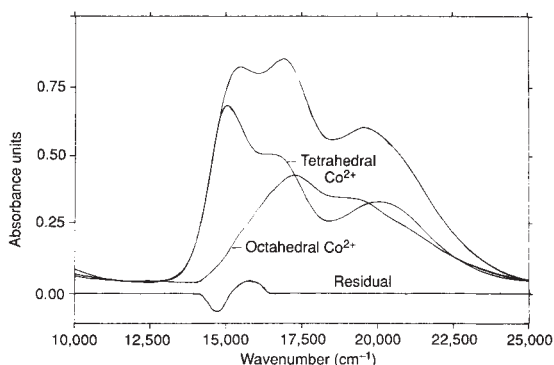


FIG. 2 Deconvolution of the spectrum of Co-doped albite glass (quenched from 100 kbar) into components due to tetrahedral and octahedral Co^{2+} .

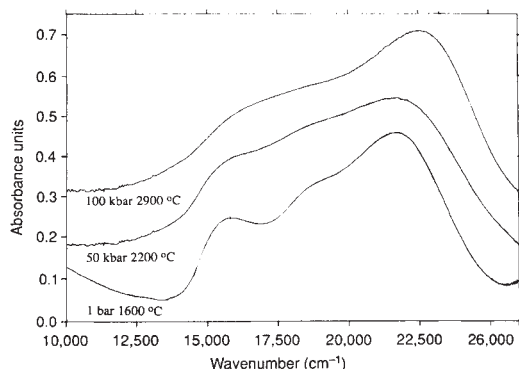


FIG. 3 Crystal field spectra of Ni-doped albite glasses quenched from different pressures and temperatures (1 wt% NiO, renormalized to 0.3 mm thickness), obtained using the same apparatus as Fig 1.

contamination by Re or MgO. Infrared spectra showed that the water content of the glasses was only about 1,000 p.p.m. by weight.

Figure 1 shows the crystal field spectra of the Co-doped glasses. The 1-atm spectrum shows the typical absorption features of tetrahedral Co^{2+} , with a system of three bands at about 15,000, 17,000, and 20,000 cm^{-1} (refs 13,14,16). Octahedral Co^{2+} would only absorb between 17,000 and 20,000 cm^{-1} (ref. 16). As pressure increases, the absorbance around 15,000 cm^{-1} , where only tetrahedral Co^{2+} absorbs, decreases, while the intensity around 20,000 cm^{-1} , where bands of both tetrahedral and octahedral Co^{2+} can occur, remains almost constant. This suggests that a significant fraction of the Co^{2+} , which was in tetrahedral coordination at 1 atm, became octahedrally coordinated at high pressure. The observed spectral changes are analogous to those observed in aqueous solutions of CoCl_2 on variation of chloride concentration, which have always been attributed to an equilibrium between four and six-coordinated Co (ref. 17). The spectra of both the Co-doped glasses and those of aqueous CoCl_2 solutions show a weak and broad band in the near-infrared (between 5,000 and 8,000 cm^{-1} in the glasses, not shown in Fig. 1) which changes only slightly on change of coordination.

Figure 2 shows a quantitative deconvolution into two components of the spectrum of the sample quenched from 100 kbar. The first component, assigned to tetrahedral Co^{2+} , is identical with the spectrum of the glass quenched at atmospheric pressure. If this spectrum is multiplied by a scaling factor of 0.5 and subtracted from the 100-kbar data, a component is obtained which is very similar to the spectrum of octahedral Co^{2+} in various reference compounds, such as the $\text{Co}(\text{H}_2\text{O})_6^{2+}$ complex in aqueous solutions¹⁸ or Co^{2+} substituted in the octahedral site of calcite (CaCO_3) (ref. 19). The deconvolution shown in Fig. 2 therefore strongly supports our interpretation that at high pressure, part of the Co^{2+} moves into an octahedral site. From the magnitude of the tetrahedral component in the 100 kbar spectrum, one can estimate that about 50% of the Co^{2+} changed coordination. This estimate is valid if the extinction coefficient of the tetrahedral species itself does not strongly depend on subtle changes in bond lengths and bond angles. Indeed, it has been shown that in sodium silicate glasses the extinction coefficient of tetrahedral Co^{2+} is almost independent of glass composition¹⁴. If the tetrahedral and octahedral component are subtracted from the 100-kbar spectrum, there is only a small residual left around 15,000 cm^{-1} . The reason for this residual is that the peak of the tetrahedral species around 15,000 cm^{-1} moves to slightly higher wavenumbers in the sample quenched from high pressure. If one allowed this peak to shift by about 500 cm^{-1} in the spectrum of the tetrahedral component, the residual would disappear entirely.

The most conspicuous change in the spectra of the Ni-doped glasses with pressure (Fig. 3) is the disappearance of the weak band around 16,000 cm^{-1} , which sometimes has been attributed to a small amount of tetrahedral Ni^{2+} (refs 12,15). If this interpretation is correct, these spectra again indicate a pressure-induced coordination change. Alternatively, this band might be due to a spin-forbidden transition of octahedral Ni^{2+} (${}^3\text{A}_{2g}(\text{F}) \rightarrow {}^1\text{E}_g(\text{D})$), enhanced by distortion¹³. At all pressures studied, however, most of the Ni^{2+} is in a distorted octahedral site, and the spectra indicate that significant changes occur at this site as well. The main absorption peak of this species shifts from 21,650 cm^{-1} (for the glass prepared at 1 atm) to 22,680 cm^{-1} in the spectrum of the glass quenched from 100 kbar. Two additional weak bands in the near infrared (not shown in Fig. 3) due to octahedral Ni^{2+} show a similar effect. The band at 9,100 cm^{-1} in the 1 atm spectrum shifts to 9,900 cm^{-1} in the 100 kbar sample. The band around 5,500 cm^{-1} also seems to shift to slightly higher frequencies with pressure.

Our spectroscopic data provide strong evidence for a pressure-induced coordination change of Co^{2+} , and possibly

Ni^{2+} , in silicate glasses from a tetrahedral to an octahedral environment. In addition, the band positions of octahedral Ni^{2+} move, indicating an increased electrostatic field strength around the Ni ion at high pressure. All of these structural changes are quenchable and therefore not due to simple elastic deformations.

A change of Co^{2+} from tetrahedral to octahedral coordination in a melt will increase the CFSE of this ion by about 29 kJ mol^{-1} ; for Ni^{2+} , the increase in CFSE is 61 kJ mol^{-1} (ref. 13). This effect will stabilize both Co^{2+} and Ni^{2+} in silicate melts relative to crystalline silicates as well as relative to a metal phase. Crystal-silicate melt and metal-silicate melt partition coefficients will be reduced by 1 (for Co) to 2 (for Ni) orders of magnitude at 1,500 K, if a complete change from four to six-fold coordination occurs. Our data suggest that this will hold true for Co^{2+} at about 200 kbar.

This effect will not be quite as large for Ni^{2+} , because in most melt compositions only a small fraction of this species is in tetrahedral sites at low pressures. A similar effect, however, will be caused by the observed increase in field strength with pressure around the octahedrally coordinated Ni^{2+} . At high pressures, Co^{2+} and Ni^{2+} will therefore behave in a much less compatible and much less siderophile way than at low pressures. This prediction is consistent with the few available experimental data on olivine/melt partitioning at high pressures²⁰. It is likely that other transition-metal ions will show similar behaviour.

This means that pressure-induced coordination changes in silicate melts have to be taken into account when modelling the global chemical evolution of the Earth and the terrestrial planets Mercury, Venus and Mars. □

Received 12 November 1992; accepted 10 May 1993.

1. Irving, A. J. *Geochim. cosmochim. Acta* **42**, 743–770 (1978).
2. Hart, S. R. & Davis, K. E. *Earth planet. Sci. Lett.* **40**, 203–219 (1978).
3. Kinzler, R. J., Grove, T. L. & Recca, S. I. *Geochim. cosmochim. Acta* **54**, 1255–1265 (1990).
4. O'Neill, H. St. C. *Geochim. cosmochim. Acta* **55**, 1159–1172 (1991).
5. Ringwood, A. E. Z. *Naturforsch.* **44a**, 891–923 (1989).
6. Drake, M. J. Z. *Naturforsch.* **44a**, 883–890 (1989).
7. Colson, R. O. *Nature* **357**, 65–68 (1992).
8. Jones, J. H., Capobianco, C. J., Drake, M. J. & O'Neill, H. *Science* **257**, 1281–1285 (1992).
9. Agee, C. B. & Walker, D. *Earth planet. Sci. Lett.* **90**, 144–156 (1988).
10. O'Neill, H. St. C. *Geochim. cosmochim. Acta* **55**, 1135–1157 (1991).
11. Seifert, S., O'Neill, H. St. C. & Brey, G. *Geochim. cosmochim. Acta* **52**, 603–616 (1988).
12. Burns, R. G. & Fyfe, W. S. *Science* **144**, 1001–1003 (1964).
13. Keppeler, H. *Am. Miner.* **77**, 62–75 (1992).
14. Nelson, C. & White, W. B. J. *Mater. Res.* **1**, 130–138 (1986).
15. Galois, L. & Calas, G. *Am. Miner.* **77**, 677–680 (1992).
16. Figgis, B. N. *Introduction to Ligand Fields* (Interscience, New York, 1966).
17. Susak, N. J. & Crerar, D. A. *Geochim. cosmochim. Acta* **49**, 555–564 (1985).
18. Greenwood, N. N. & Earnshaw, A. *Chemistry of the Elements* (Pergamon, Oxford, 1984).
19. Rossman, G. R. in *Spectroscopic Methods in Mineralogy and Geology* (ed. Hawthorne, F. C.) 207–254 (Mineralogical Society of America, Washington, 1988).
20. McFarlane, E. A. & Drake, M. J. in *Origin of the Earth* (eds Newsome, H. E. & Jones, J. H.) 135–150 (Oxford, New York, 1990).

A strategy of win–stay, lose–shift that outperforms tit-for-tat in the Prisoner's Dilemma game

Martin Nowak* & Karl Sigmund†

*Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK

†Institut für Mathematik, Universität Wien, Strudlhofgasse 4, A-1090 Vienna, Austria

THE Prisoner's Dilemma is the leading metaphor for the evolution of cooperative behaviour in populations of selfish agents, especially since the well-known computer tournaments of Axelrod¹ and their application to biological communities^{2,3}. In Axelrod's simulations, the simple strategy tit-for-tat did outstandingly well and subsequently became the major paradigm for reciprocal altruism^{4,12}. Here we present extended evolutionary simulations of heterogeneous ensembles of probabilistic strategies including mutation and selection, and report the unexpected success of another protagonist: Pavlov. This strategy is as simple as tit-for-tat and embodies the fundamental behavioural mechanism win–stay, lose–shift, which seems to be a widespread rule¹³. Pavlov's success is based on two important advantages over tit-for-tat: it can correct occasional mistakes and exploit unconditional cooperators. This second feature prevents Pavlov populations from being undermined by unconditional cooperators, which in turn invite defectors. Pavlov seems to be more robust than tit-for-tat, suggesting that cooperative behaviour in natural situations may often be based on win–stay, lose–shift.

Two players engaged in the Prisoner's Dilemma have to choose between cooperation (C) and defection (D). In any given round, the two players receive R points if both cooperate and only P points if both defect; but a defector exploiting a cooperator gets T points, while the cooperator receives S (with $T > R > P > S$ and $2R > T + S$). Thus in a single round it is always best to defect, but cooperation may be rewarded in an iterated (or spatial¹⁴) Prisoner's Dilemma.

The conspicuous success of the tit-for-tat (TFT) strategy (start with a C, and then use your co-players previous move) relies in part on the clinical neatness of a deterministic cyber-world. In natural populations, errors occur^{7,12}. TFT suffers from stochastic perturbations in two ways: (1) a TFT population can be 'softened up' by random drift introducing unconditional cooperators, which allow exploiters to grow (TFT is not an evolutionarily stable strategy^{15,16}); and (2) occasional mistakes between two TFT players cause long runs of mutual backbiting. (Such mistakes abound in real life: even humans are apt to vent frustrations upon innocent bystanders.)

Within the restricted world of strategies reacting only to the co-players previous move, TFT has a very important, but transitory role: in small clusters, it can invade populations of defectors, but then bows out to a related strategy, 'generous tit for tat' (GTFT), which cooperates after a co-player's C, but also with a certain probability after a D⁹.

But as soon as one admits strategies which take into account the moves of both players in the previous round, evolution becomes much less transparent¹⁷. We first conjectured that GTFT (or variants thereof) would win the day, but are forced to admit, after extensive simulations, that the strategy Pavlov did much better in the long run. A Pavlov player cooperates if and only if both players opted for the same alternative in the previous move. The name¹⁸ stems from the fact that this strategy embodies an almost reflex-like response to the payoff: it repeats its former move if it was rewarded by R or T points, but switches behaviour if it was punished by receiving only P or S points. This strategy, which went by the name of 'simpleton'¹⁹, fares poorly against inveterate defectors: in every second round, it switches to cooperation. It cannot gain a foothold in a defector's world; defectors have to be invaded by other strategies, like TFT⁹. But Pavlov has two important advantages over TFT: (1) an inadvertent mistake between players using TFT causes a drawn-out battle; between two Pavlovians, it causes one round of mutual defection followed by a return to joint cooperation²⁶. Thus Pavlov is fairly tolerant, like GTFT, and can correct mistakes. (2) Whereas TFT and GTFT can be invaded by drift by all-out cooperators (to the eventual profit of exploiters), Pavlov has no qualms in exploiting a sucker, once it has discovered (after an accidental mistake) that it need not fear any retaliation.

Softies cannot subvert a Pavlov population. In this sense, cooperation based on Pavlov is a safer bet than TFT. Pavlov's advantage shows best among nice strategies.

For our simulations, we consider all (stochastic) strategies with memory one (that is, recalling only the previous round). These strategies are defined by the conditional probabilities (p_1, p_2, p_3, p_4) to cooperate, given the outcome of the previous round was R, S, T or P , respectively. The game between two such strategies can be formulated as a Markov process, and its stationary distribution specifies the payoff for the infinitely iterated Prisoner's Dilemma²⁰. Owing to noise, the initial move has no effect in the long run. For mathematical simplicity we retain the assumption of the infinitely iterated game, but note that the outcome is essentially unchanged if we consider sufficiently long iterated games. In our notation TFT is given by (1, 0, 1, 0) and Pavlov by (1, 0, 0, 1), but mistakes in implementing the move change 1 to $1 - \varepsilon$ and 0 to ε , where ε is a small number specifying the minimal amount of noise. This is closely related to the 'trembling hand' in Selten's game theoretical notion on perfect equilibrium^{15,21}.

We start each simulation with the random strategy (0.5, 0.5, 0.5, 0.5). Every 100 generations (on average), we introduced a small amount of a randomly chosen mutant strategy. the fre-

quencies of strategies spread according to the usual game dynamics^{22,23}, reflecting natural selection. Strategies with higher payoffs produce more offspring. Strategies whose frequency dropped below a certain threshold were discarded. Each run was observed for 10^7 generations, generating a total of about 10^5 different mutant strategies. (Note that the timescale is arbitrary, because the difference equation can be seen as an approximation of a differential equation. It is, however, very important to study the long-term dynamics and to try many mutants.) The evolutionary chronicles display an extreme diversity. Nevertheless, they allowed some clear and simple conclusions: (1) the plot for the average payoff in the population is a show-piece of punctuated equilibria (Fig. 1). Most of the time, this payoff is very close to one of the extremal values P (a regime of defection) or R (overall cooperation). The time for switching from one of the regimes to the other is usually extremely short (only a few generations). The periods of stasis frequently last for millions of generations. The later in the run, the longer they last. But the threat of a sudden collapse may never abate. (2) There is a clear tendency for cooperation (Fig. 2). After $t=10^4$ generations, only 27.5% of the runs exhibit cooperation (population payoff > 2.95); but 90% at $t=10^7$. There is also a clear tendency towards Pavlov, which dominates 10% of the runs at $t=10^4$, but

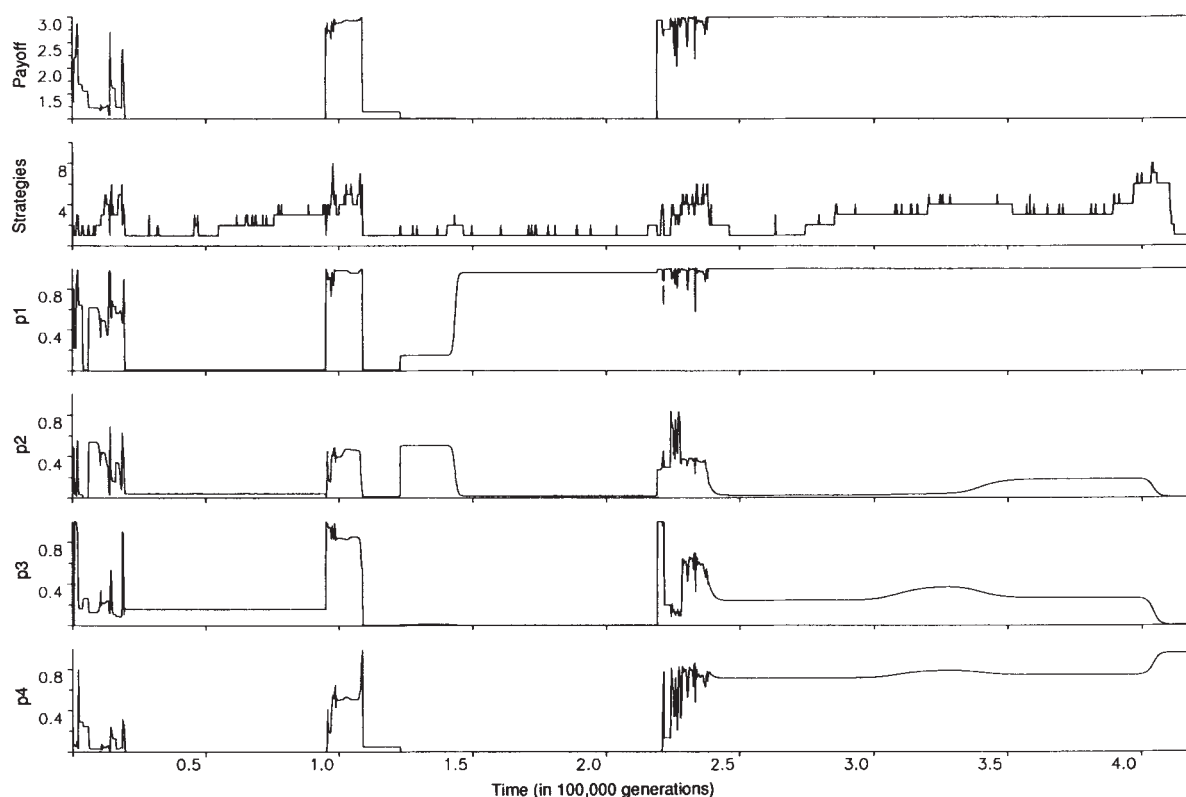


FIG. 1 An evolutionary simulation (including mutation and selection) of all strategies that consider the previous move in the iterated Prisoner's Dilemma. Such strategies are defined by four probabilities (p_1, p_2, p_3, p_4) that cooperate after having received payoff R, S, T or P in the previous round. We start with the random strategy (0.5, 0.5, 0.5, 0.5). In each generation there is a 0.01 probability that a new mutant strategy is generated at random. For technical reasons, we used the U-shaped density distribution $[\pi x(1-x)]^{-1/2}$ for sampling the p_i in the interval 0, 1. This distribution, which is familiar to geneticists, helps to explore the corners of the four-dimensional cube of stochastic strategies in a more efficient way. We assume that there is a minimal amount of noise; thus we restrict $0.001 < p_i < 0.999$. The frequencies of strategies grow according to the usual game dynamics. Strategies with frequencies below 0.001 are removed. In this example, the initial struggle for cooperation is unsuccessful. An *AIID* like strategy emerges as winner

and dominates until $t \approx 92,000$. Then a TFT mutant invades and establishes a regime of cooperation dominated by GTFT. This population is undermined by more and more forgiving strategies (increasing p_4), which leads to a breakdown of cooperation and another period of defection, now dominated by the severe retaliator 'GRIM' (0.999, 0.001, 0.001, 0.001). Again TFT invades and catalyses the rise of cooperation. After some small adjustments in p_2, p_3 , and p_4 , there is a long lasting period of cooperation dominated by the Pavlov-like strategy (0.999, 0.001, 0.007, 0.946). This persists at least until $t=10^7$ (not shown). The figure shows the average population payoff, the number of strategies present at a given time, and the population averages of the probabilities p_1, p_2, p_3 , and p_4 . (The time is given in generations of the difference equation, but the timescale is arbitrary if the difference equation is understood as an approximation of a differential equation.) Payoff values: $R=3, S=0, T=5, P=1$.

82.5% at $t=10^7$. Only a few runs are eventually dominated by GTFT-like behaviour. (3) Usually, the population is monomorphic or very mildly polymorphic. We rarely find more than 10 strategies in one population. Cooperative populations (with an average population payoff close to R) are normally dominated by one or two strategies very close to Pavlov, although some long hegemonies by larger mixtures of GTFT-like strategies can also be observed. (4) Pavlov can be invaded by *AllD* if $2R < T + P$, but a prudent variant of it, $(1, 0, 0, x)$ with $x < (R - P)/(T - R)$, is stable against *AllD*. For $2R = T + P$ (as is the case with Axelrod's payoff values), Pavlov can be invaded by *AllD*, but the stochastic, Pavlov-like strategy $(0.999, 0.001, 0.001, 0.995)$ cannot. We observe that strategies very close to this 'almost Pavlov' win most evolutionary runs for $R = 3$. Figure 3 shows the outcome of computer simulations for various different values of R . For most values of R , Pavlov clearly dominates.

How does this relate to real biology? It might well be that many cases of cooperation based on reciprocal altruism are due to Pavlov rather than to TFT. In more natural set-ups, retaliation has usually been interpreted as evidence for TFT^{4-6,27}. But these experiments seem to be consistent with a Pavlov-like strategy as well. It would speak for Pavlov if animals have a tendency to exploit non-retaliators, or if they are apt to resume cooperation after bilateral defection. (In a sense, we observe Pavlov-type behaviour daily ourselves. Usually, a domestic misunderstanding causes a quarrel, after which cooperation is resumed; and the advice 'never to give a sucker an even break' is frequently adopted among members of our species.) Incidentally, in natural encounters where different rounds are not clearly separated, there is a problem in 'timing' the recovery—we may expect signals to evolve as cues.

The simple learning rule, win-stay, lose-shift, seems to be widespread and works in many other contexts. It is a particular instance of the 'law of effect', which states that animals perform more rewarded behaviours and less non-rewarded behaviours¹³. It is plausible that more sophisticated variants become established among higher animals, which stick to a move as long as a weighted payoff from the last few rounds is sufficiently rewarding (in the manner of Harley's learning rule²⁴).

By and large, the iterated Prisoner's Dilemma has been seen as a story of TFT, but our results suggest that cooperation based on win-stay, lose-shift may be more robust. The success of

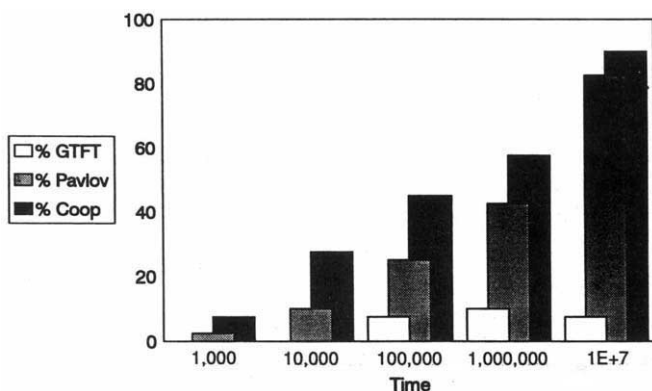


FIG. 2 There is a tendency towards cooperation and Pavlov. The figure shows the percentage of populations (at various times, arbitrary units) that are dominated by GTFT, Pavlov, or cooperative behaviour in general (population payoff average > 2.95). Note that later in the runs cooperation is always based on either Pavlov or GTFT-like behaviour. 40 simulations were performed. In each simulation $\sim 10^5$ different mutant strategies were generated at random. For each strategy the conditional probabilities to cooperate were within 0.001 and 0.999, thus the minimal amount of noise was $\epsilon = 0.001$. Pavlov-like behaviour is characterized by strategies close to $(0.999, 0.001, 0.001, 0.995)$, and GTFT by strategies around $(0.999, 0.33, 0.999, 0.33)$. Payoff values: $R = 3$, $S = 0$, $T = 5$, $P = 1$ (as in Axelrod's simulations).

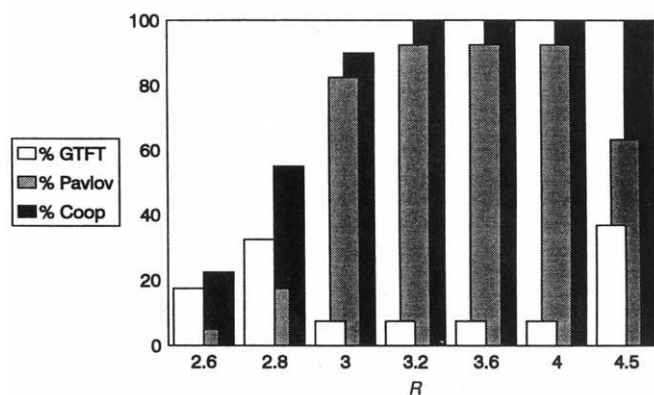


FIG. 3 Pavlov clearly outperforms TFT-like behaviour for a variety of different payoff values, R . According to the rules, R can vary between $(T+S)/2$ and T . For low values of R , GTFT seems to be slightly more successful than Pavlov, but in this case it is very difficult for cooperation to get established at all. Here the Prisoner's Dilemma is obscured by the fact that alternating C and D is almost as rewarding as cooperation. For high values of R , the populations are quite often dominated by GTFT, which for this parameter region is very close to *AllC*. For each value of R we performed 40 simulations. Pavlov-like behaviour is characterized by strategies around $(0.999, 0.001, 0.001, x)$, where $x = 0.999$ for $R > 3$, $x \approx 0.995$ for $R = 3$, and $x = (R - P)/(T - R)$ for $R < 3$. GTFT-like behaviour is characterized by strategies around $(0.999, y, 0.999, y)$, where $y = \min\{1 - (T - R)/(R - S), (R - P)/(T - P)\}$. Payoff values: $S = 0$, $T = 5$, $P = 1$ and R as indicated.

Pavlov-like behaviour does not seem to be restricted to strategies, which only remember the last move. In other evolutionary runs, where mutations can extend the memory length, similar strategies have been found: typically, they resume cooperation after two rounds of mutual defection^{25,26}. Like Pavlov, they correct for mistakes and exploit unconditional cooperators. There is, of course, no limit in the complexity of conceivable strategies. But it may be expected that the simple, natural rule of win-stay, lose-shift performs well under a variety of sophisticated conditions. Pavlov is no simpleton. □

Received 1 February; accepted 15 April 1993.

1. Axelrod, R. *The Evolution of Cooperation* (Basic Books, New York, 1984).
2. Axelrod, R. & Hamilton, W. D. *Science* **211**, 1390-1396 (1981).
3. Axelrod, R. & Dion, D. *Science* **242**, 1385-1390 (1988).
4. Wilkinson, G. *Nature* **306**, 181-184 (1984).
5. Lombardo, M. P. *Science* **227**, 1363-1365 (1985).
6. Milinski, M. *Nature* **325**, 433-435 (1987).
7. May, R. M. *Nature* **327**, 15-17 (1987).
8. Dugatkin, L. A. *Behav. Ecol. Sociobiol.* **25**, 395-397 (1988).
9. Nowak, M. & Sigmund, K. *Nature* **355**, 250-253 (1992).
10. Krebs, J. R. & Davies, N. B. *An Introduction to Behavioural Ecology* (Sinauer, MA, 1981).
11. Dawkins, R. *The Selfish Gene* (Oxford Univ. Press, Oxford, 1988).
12. Sigmund, K. *Games of Life* (Oxford Univ. Press, Oxford, 1993).
13. Domjan, M. & Burkhard, B. *The Principles of Learning and Behaviour* (Brooks/Cole, Monterey, 1986).
14. Nowak, M. A. & May, R. M. *Nature* **359**, 826-829 (1992).
15. Selten, R. & Hammerstein, P. *Th. Behav. Brain Sci.* **7**, 115-142 (1984).
16. Boyd, R. & Lorbbaum, J. P. *Nature* **327**, 58-59 (1987).
17. Nowak, M. & Sigmund, K. *Proc. natn. Acad. Sci. U.S.A.* **90**, 5091-5094 (1993).
18. Kraines, D. & Kraines, V. *Theory and Decision* **26**, 47-63 (1988).
19. Rapoport, A. & Chammah, A. M. *Prisoner's Dilemma* (Univ. of Michigan Press, Ann Arbor, 1965).
20. Nowak, M. & Sigmund, K. *Acta appl. Math.* **20**, 247-265 (1990).
21. Boyd, R. *J. theor. Biol.* **136**, 47-56 (1989).
22. Maynard Smith, J. *Evolution and the Theory of Games* (Cambridge Univ. Press, Cambridge, 1982).
23. Hofbauer, J. & Sigmund, K. *The Theory of Evolution and Dynamical Systems* (Cambridge Univ. Press, Cambridge, 1988).
24. Maynard Smith, J. *Th. Behav. Brain Sci.* **7**, 95-101 (1984).
25. Axelrod, R. in *Genetic Algorithms and Simulated Annealing* (ed. Davis, D.) (Pitman, London, 1987).
26. Lindgren, K. in *Artificial Life II* (eds Farmer, D. et al.) (Proc. Santa Fe Inst. Stud., Addison-Wesley, 1991).
27. Reboreda, J. C. & Kacelnik, A. *J. exp. Animal Behav.* **60**, 176-193 (1993).

ACKNOWLEDGEMENTS We thank W. Hamilton, A. Kacelnik, J. Krebs and R. May for discussion.

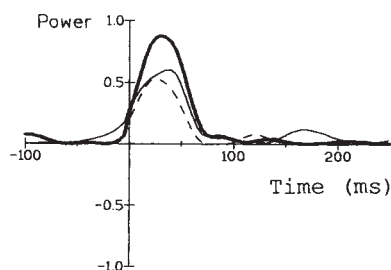
Selective attention enhances the auditory 40-Hz transient response in humans

H. Tiittinen, J. Sinkkonen, K. Reinikainen, K. Alho, J. Lavikainen & R. Näätänen

Cognitive Psychophysiology Research Unit, Department of Psychology, University of Helsinki, Ritarikatu 5, SF-00170 Helsinki, Finland

STUDIES of human auditory¹⁻³ and somatosensory³ modalities have shown that there is an oscillatory response in the γ -band (at about 40 Hz) frequency which is elicited by either steady state¹⁻³ or transient⁴ stimulation. The auditory 40-Hz response is generated at least partially in the auditory cortex^{4,5} as a result of thalamo-cortical interaction⁶ and may serve perceptual integration^{7,8} and conscious perception⁹. A connection to selective attention has been implied in human¹⁰ and animal¹¹ studies, although the evidence is inconclusive¹². Moreover, fundamental differences between the human and animal 40-Hz responses¹³ prohibit generalization. Furthermore, most experiments have used steady-state stimulation during which the brain does not regain its resting state between stimuli as it does when transient stimulation is used¹⁴. Here we study the effect of selective attention on the auditory γ -band (40-Hz) transient response using subjects listening to tone pips presented in one ear while ignoring a concurrent sequence of tone pips in the other ear. The 40-Hz response was larger when subjects paid attention to stimuli rather than ignored them. This attention effect was most pronounced over the frontal and central scalp areas. Our results demonstrate a physiological correlate of selective attention in the 40-Hz transient response in humans.

We studied 40-Hz activity in a dichotic-listening protocol in eight subjects (mean age 26 years; two females) with normal hearing. Subjects were instructed to attend to stimuli presented to a designated ear and to press a button every time they detected an occasional deviant tone among the frequent ('standard') tones in that ear and to ignore concurrent stimuli in the opposite ear ('Attend' condition). In a separate control condition subjects were reading a book while the same auditory stimulus sequences were presented ('Reading' condition). Scalp-recorded 40-Hz transient responses were averaged separately for Reading and Attend conditions and, within the latter, separately for the attended and unattended input.



Thirty-six blocks of 500 tone pips (intensity 80 dB SPL, duration 60 ms, rise and fall times 10 ms each) were presented to each subject. Each block consisted of standard (left ear 1,000 Hz, right ear 500 Hz; $P=0.475$ for each) and deviant stimuli (left 1,050 Hz, right 475 Hz; $P=0.025$ for each) presented through earphones in a random order. The interstimulus interval was constant at 700 ms. Half of the blocks were presented in reverse order.

Figure 1 presents the Gabor-filtered^{15,16} 40-Hz responses of a representative subject to the left-ear standard stimuli recorded over the frontal area. The responses to attended stimuli (thick line) were clearly larger than those obtained when the same stimuli were unattended (thin line), or when the subject was reading a book (dashed line).

In Fig. 2, the response of a representative subject to left-ear standard stimuli is presented as a three-dimensional time-frequency map, with power represented on the vertical axis. The map reveals that shortly after stimulus onset, the activity in the 30–60-Hz range is distributed around a distinct peak at 40 Hz and, further, that attention affects the power landscape by enhancing this peak.

Figure 3 shows the grand-average 40-Hz transient responses to standard stimuli. Responses to attended stimuli were enhanced, especially in the frontal sites, compared with responses to unattended stimuli and to those elicited in the Reading condition. The 40-Hz response peaked at around 25 ms from stimulus onset and lasted for about 100 ms. Response to 1,000-Hz stimuli were slightly more pronounced and longer than the corresponding responses to 500-Hz stimuli. The 40-Hz response was statistically significant at Fz and Cz (see Fig. 1 legend) for both ears (Wilcoxon test; $T \leq 1$, $P < 0.008$, one-tailed). Two-way ANOVA (Friedman nonparametric; factors: subject, condition; mean power during 0–100 ms) revealed a significant attention effect at Fz and Cz ($F=7.75$, d.f. = 2, $P < 0.025$). The 40-Hz response to attended stimuli at Fz was always significantly larger than to unattended stimuli and in the Reading condition (one-tailed Wilcoxon tests for pairs; $T=2$, $P < 0.012$, and $T=1$, $P < 0.008$, respectively).

The effect of attention was also tested at lower frequencies (passbands of 0–30, 0–20 and 10–30 Hz; mean amplitude during either 0–50 or 20–50 ms from stimulus onset). All the analyses comparing the responses to attended and unattended stimuli proved to be statistically nonsignificant, which is further evidence that the attention effect can only be found at or near the 40-Hz range.

FIG. 1 40-Hz responses (from electrode Fz) of a representative subject to left-ear 1,000-Hz standard stimuli when attended (thick line), unattended (thin line), and when the subject was reading (dashed line). Selective attention enhances the response to attended stimuli. The EEG was recorded (passband 0.1–100 Hz, sampling rate 200 Hz) from the scalp midline (electrodes Fpz, Fz, Cz, Pz and Oz of the international 10–20 system) and from the canthus of the right eye (Heog) with Ag–AgCl electrodes referenced to linked mastoids. The EEG was averaged using an analysis period of 600 ms, including a 200-ms prestimulus baseline. Artefacts were rejected by automatically discarding epochs with absolute value of the amplitude exceeding 75 μ V at any electrode. The EEG was digitally convoluted by a Gabor wavelet^{15,16}, yielding a continuous measure of frequency-specific power over time. The Power measure $A(t)$ is the squared norm of a (complex) Gabor-filtered averaged response, that is $A(t) = |h * y(t)|^2$, where the asterisk signifies convolution, $y(t)$ is the signal, and $h(t)$ is the Gabor wavelet: $h(t) = \exp(-x^2/2 + imx)$, $x = 2\pi \times 40\text{Hz} \times t/m$, $m=5.3$, and i is the imaginary unit. The gaussian frequency gain function is centred at 40 Hz, s.d. = 8 Hz. The level of the EEG noise was estimated as the mean power during the period between 100–50 ms preceding stimulus onset, and this was subtracted from the signal before further analyses. (Note that the power of a limited frequency band cannot be measured without temporal spreading of the response that gives rise to an apparent response before stimulus onset; see Fig. 3.)

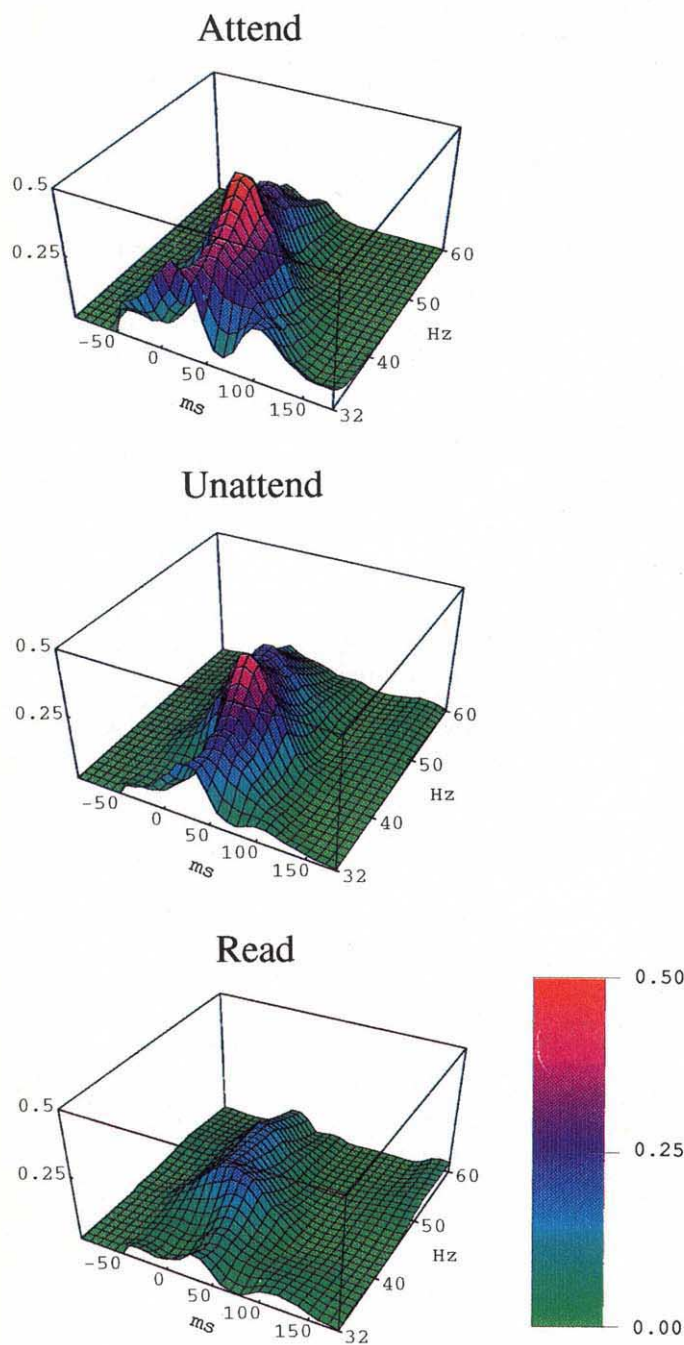


FIG. 2 The response of a representative subject to 1,000-Hz left-ear standard stimuli presented as a three-dimensional time-frequency map. Power, represented on the vertical axis, is measured by Gabor wavelets with $m=7.5$, and has been linearly interpolated between discrete frequencies. Frequencies are normalized by setting the sum of the squared norms of the convolution coefficients, $\sum |h(t)|^2$, constant for all frequencies.

In conclusion, our results show that selective attention enhances the amplitude of the auditory 40-Hz response and hence demonstrate a physiological correlate of selective attention in this response. It remains to be determined whether this effect, observed in averaged data, originates from a diminished time jitter between consecutive single responses or their increased amplitude. As the response to unattended stimuli was in most cases (Fig. 3) at least as large in amplitude as that in the Reading

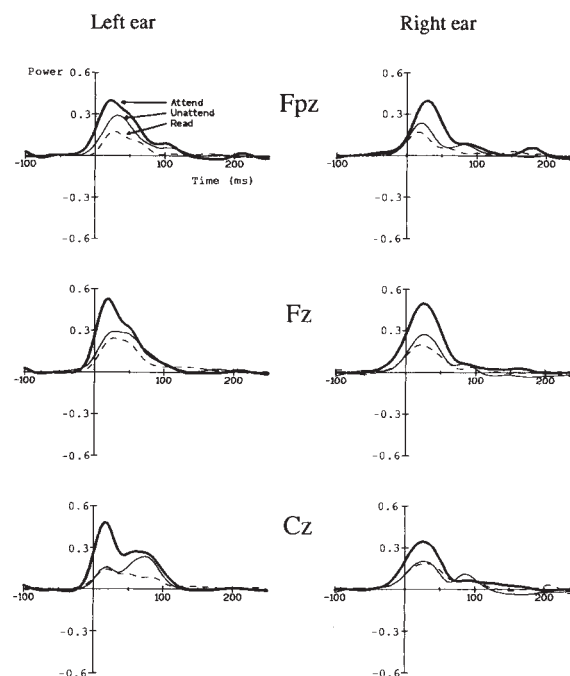


FIG. 3 Gabor-filtered grand-average 40-Hz responses measured from the scalp at Fpz, Fz and Cz to left- (1,000 Hz) and right-ear (500 Hz) standard stimuli when attended (thick line), when stimuli in the opposite ear were attended (thin line), or when the subject was reading a book (dashed line).

condition, it could be concluded that, in contrast to many leading views¹⁷⁻²⁰, the attentional mechanism reflected by the 40-Hz response is not suppressing (for example, gating or 'filtering') the processing of unattended stimuli. It has also been proposed^{6,21,22} that the 40-Hz response could be used for diagnostic purposes. As the attentional enhancement of the 40-Hz response is reliable even at the single-subject level, it might indeed prove to be a powerful tool for studying attentional disorders in clinical groups such as schizophrenics and patients with Alzheimer's disease. □

Received 1 September 1992; accepted 8 April 1993.

- Galambos, R., Makeig, S. & Talmachoff, P. J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2643-2647 (1981).
- Makeig, S. & Galambos, R. *Soc. Neurosci. Abstr.* **15**, 113 (1989).
- Galambos, R. *Ann. N.Y. Acad. Sci. U.S.A.* **388**, 722-728 (1982).
- Pantev, C. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 8996-9000 (1991).
- Mäkelä, J. P. & Hari, R. *Electroenceph. clin. Neurophys.* **66**, 539-546 (1987).
- Ribary, U. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 11037-11041 (1991).
- Gray, C. M. et al. *Nature* **338**, 334-337 (1989).
- Engel, A. K. et al. *Trends Neurosci.* **15**, 218-226 (1992).
- Llinás, R. & Pare, D. *Neuroscience* **44**, 521-535 (1991).
- Sheer, D. E. In *Self-regulation of the Brain and Behaviour* (eds Elbert, T., Rockstroh, B., Lutzenberger, W. & Birbaumer, N., 64-84 (Springer, Berlin, 1984).
- Rougeul, A. et al. *Electroenceph. clin. Neurophys.* **46**, 310-319 (1979).
- Linden, R. D. et al. *Electroenceph. clin. Neurophys.* **66**, 145-159 (1987).
- Galambos, R. in *Induced Rhythms in the Brain*, (eds Bullock, T. H. & Başar, E., 201-216 (Birkhäuser, Boston, 1991).
- Kaufman, L. & Williamson, S. J. in *Auditory Evoked Magnetic Fields and Electric Potentials* (eds Grandori, F., Hoke, M. & Romani, G. L.) 283-312 (Karger, Basel, 1990).
- Daubechies, I. *IEEE Transactions on Information Theory* **36**, 961-1005 (1990).
- Gabor, D. *Proc. IEE* **93**, 429-457 (1946).
- Harter, M. R. & Aine, C. J. in *Varieties of Attention* (eds Parasuraman, R. & Davies, D. R. Academic, London (1984).
- Skinner, J. E. & Yingling, C. D. *Progr. clin. Neurophys.* **1**, 30-69 (1977).
- Hillyard, S. A. & Mangun, G. R. in *Current Trends in Event-Related Potential Research EEG suppl.* **40**, (eds Johnson, R., Parasuraman, R. & Rohrbaugh, J. W.) 61-67 (Elsevier, Amsterdam 1984).
- Oatman, L. *Exp. Neurol.* **32**, 341-356 (1971).
- Ribary, U. et al. *Eur. J. Neurosci. Suppl.* **1**, 44.17 (1988).
- Ribary, U. et al. in *Advances in Biomagnetism* (eds Williamson, S., Hoke, M., Stroink, G. & Kotani, M.) 311-314 (Plenum, New York, 1989).

ACKNOWLEDGEMENTS. We thank P. May, P. Paavilainen, W. Teder and V. Virsu for their constructive comments on the manuscript. This study was supported by the Academy of Finland, the Finnish Cultural Foundation, and the Emil Aaltonen Foundation (Tampere, Finland).

Role of H5 domain in determining pore diameter and ion permeation through cyclic nucleotide-gated channels

Evan H. Goulding*, Gareth R. Tibbs†, David Liu & Steven A. Siegelbaum††

Center for Neurobiology and Behavior, Departments of * Physiology and † Pharmacology, † Howard Hughes Medical Institute, Columbia University, 722 West 168 Street, New York, New York 10032, USA

ION permeation through membrane channels is thought to be governed by a narrow region of the channel pore termed the selectivity filter¹, which has been proposed to discriminate among ions by both specific binding and molecular sieving, as determined by pore diameter. Recent evidence suggests that a conserved domain (known as H5, P or SS1-SS2) in voltage-gated potassium²⁻⁸, sodium⁹⁻¹³ and calcium¹² channels contributes to the lining of the pore. Here we investigate whether the H5 domain determines pore diameter and examine the role of pore diameter in controlling ion permeation. These studies rely on differences in single channel conductance, ion selectivity and apparent pore diameter between cyclic nucleotide-gated channels cloned from bovine retina¹⁴ and catfish olfactory neurons¹⁵. Using chimaeric retinal-olfactory channels, we find that the H5 domain determines these differences in permeation properties, providing structural evidence that the cyclic nucleotide-gated channels are indeed members of the voltage-gated channel family¹⁵⁻¹⁷. Moreover, these results show directly that the H5 domain helps form the selectivity filter and that molecular sieving is important in controlling ion permeation.

As shown in Figs 1 and 2, there is a marked difference in conductance between retinal and olfactory cyclic nucleotide-gated channels expressed in *Xenopus* oocytes. At positive membrane potentials, both channels exhibit a single open conductance state as shown by the single open peak in current amplitude histograms, with slope conductances of 21 pS for the retinal channel and 51 pS for the olfactory channel (Fig. 2). At negative potentials, the retinal channel shows a single open state with a slope conductance of 18 pS (Fig. 2), whereas the olfactory channel displays a prominent subconductance state (32 pS) in addition to the main open state (60 pS). As a result, the amplitude histogram of the olfactory channel displays two open current peaks (Fig. 1b).

We examined the role of H5 in determining channel conductance by constructing a chimaeric channel (retinal-olfactory channel) in which the retinal channel H5 domain was replaced with the corresponding region of the olfactory channel (Fig. 1a). As shown in Figs 1 and 2, the chimaeric channel has conductance properties similar to those of the olfactory channel. Thus, at positive potentials, the chimaeric channel displays a single open state with a large unitary conductance of 68 pS. At negative membrane potentials, the channel displays a large conductance state of 66 pS and a prominent subconductance state of 39 pS. The somewhat higher conductance of the chimaeric channel compared with the olfactory channel may result from novel interactions between the olfactory H5 region and the rest of the retinal channel or may reflect a small contribution of some other region of the channel to the pore¹⁸.

The increased conductance of the chimaeric channel is not due to a nonspecific disruption of pore structure because a converse chimaera (olfactory channel containing the retinal H5 domain) displays a small conductance appropriate for the retinal channel (chord conductance of 17 pS at -80 mV, $n=2$; see legend to Fig. 2 for details). Unfortunately, the low expression of this

olfactory-retinal chimaera prevented us from studying it further. All subsequent studies were carried out on the retinal-olfactory H5 chimaera.

The effects of substituting the H5 domain are specific for ion permeation. Thus, the gating of the retinal-olfactory chimaera by cyclic GMP is identical to that of the retinal channel. Dose-response curves for the activation of the chimaeric and retinal

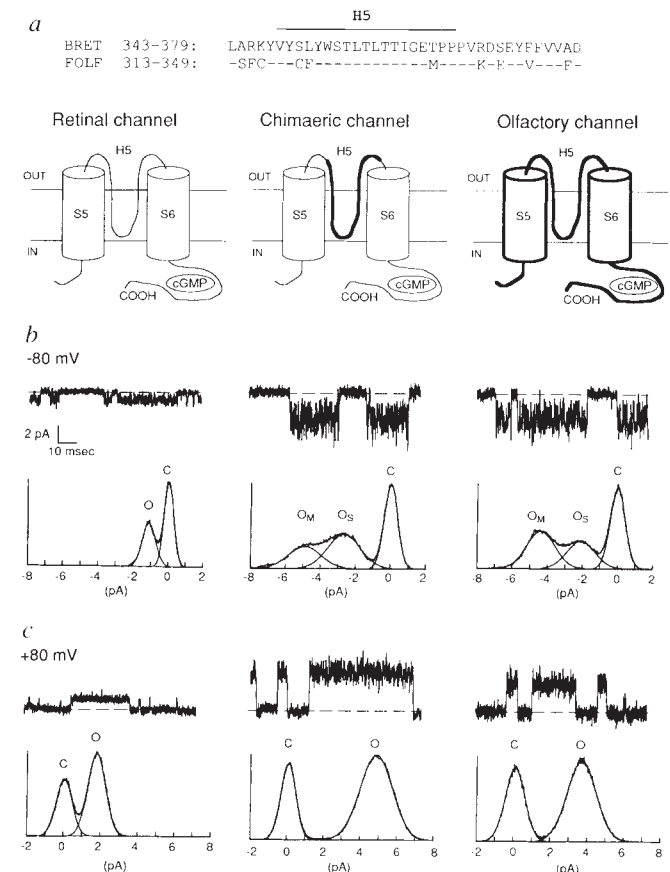
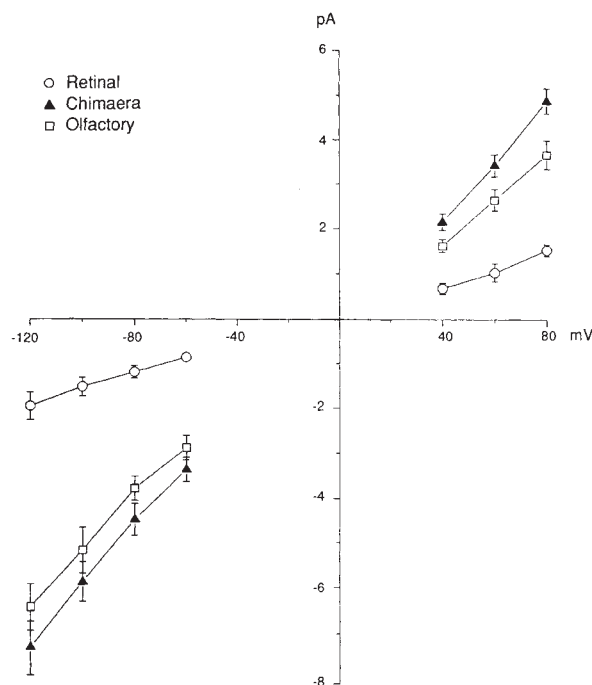


FIG. 1 Comparison of the H5 region amino-acid sequences and the single channel currents of the retinal, retinal-olfactory chimaeric, and olfactory cyclic nucleotide-gated (CNG) channels. **a**, Amino-acid sequences of H5 and its flanking regions are compared for a bovine retinal¹⁴ (BRET) and catfish olfactory CNG channel¹⁵ (FOLF). Dashes represent identical residues. In the retinal-olfactory chimaera, amino acids 343-379 from the retinal CNG channel have been replaced by amino acids 313-349 from the olfactory CNG channel. The diagram illustrates positions of these residues relative to the S5 and S6 segments. **b** and **c**, Single channel current records (top) and amplitude histograms (bottom) from representative inside-out patches for retinal, chimaeric and olfactory CNG channels at -80 mV (**b**) and $+80$ mV (**c**). Dashed lines indicate closed current level. Amplitude histograms show superimposed fits of 2 or 3 gaussian functions.

METHODS. *Xenopus* oocytes were prepared and injected with *in vitro* transcribed RNA as described¹⁵. The chimaera was made using the polymerase chain reaction (PCR). Two PCR products generated from a retinal channel template were isolated on gels and used as primers to generate a chimaeric PCR product from an olfactory channel template. The chimaeric PCR product was then digested with *SauI* and *NsiI*. The 0.9-kb insert was subcloned into a bovine retinal cDNA in pGEM-3Z (Promega) containing the 5' and 3' *Xenopus* β -globin untranslated sequences flanking the polylinker region²¹. The presence of the chimaeric channel cDNA was verified by double-stranded dideoxy chain termination sequencing of the entire fragment amplified by PCR. Single-channel currents and amplitude histograms were obtained from inside-out patches as described¹⁵. Pipette and bath solutions were identical and contained (in mM): 125 K, 97 Cl, 10 HEPES, 10 EGTA, 1 EDTA, pH 7.2 (titrated with KOH).



channels by cGMP (at -80 mV) yield $K_{1/2}$ values for cGMP of $65 \pm 19 \mu\text{M}$ ($n=4$, s.d.) and $71 \pm 25 \mu\text{M}$ ($n=15$), respectively, and identical values for the Hill coefficients of 2.0 ± 0.3 . Like the retinal channel, the chimaeric channel shows minimal activation by 1 mM cAMP ($<1\%$ of the maximal response to 1 mM cGMP), whereas the olfactory channel is activated equally by

FIG. 2 Single channel current-voltage (I - V) relationship for the retinal, chimaeric, and olfactory CNG channels. Only the main conductance states are shown for the chimaeric and olfactory channels. Current amplitudes were determined by fitting gaussian distributions to amplitude histograms as in Fig. 1. Each point is mean $\pm$ s.d. of three to nine determinations from different patches. Data were obtained as for Fig. 1. To control for possible nonspecific effects in the chimaera, we constructed converse chimaeras in which different portions of the retinal H5 domain and flanking sequences were substituted for the corresponding region of the olfactory channel. In the first construct, ORH, we substituted the entire region indicated in Fig. 1a (amino acid residues 343–379 in the retinal channel were exchanged for residues 313–349 of the olfactory channel). The second construct, ORH-B, exchanged only the putative H5 domains (overlined region of Fig. 1a; residues 349–365 in the retinal channel for residues 319–335 in the olfactory channel). ORH showed very low expression, with brief infrequent and noisy openings that showed no resolvable peaks in amplitude histograms, although most openings were <2 pA. ORH-B gave better expression, with longer well-defined openings of 1.4 pA at -80 mV, yielding a chord conductance of 17 pS ($n=2$; data not shown), similar to the 15 pS chord conductance of the retinal channel at -80 mV.

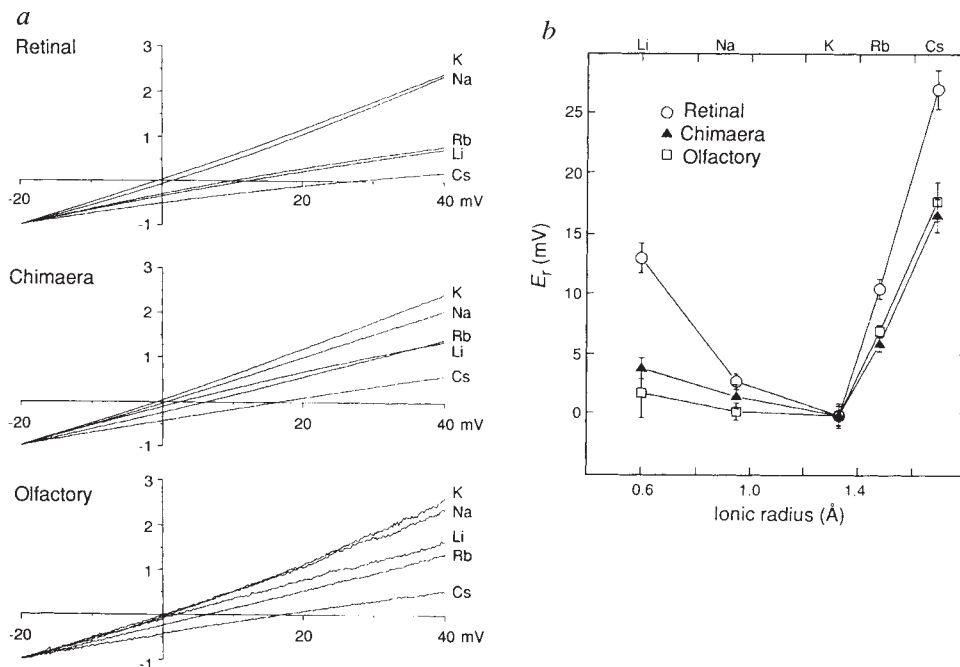
METHODS. ORH and ORH-B chimaeric inserts were made using PCR as described for Fig. 1. The chimaeric PCR products were digested with *MluI* and *Clal*. The 0.3 -kb insert was then cloned into a catfish olfactory cDNA in the PGEM-3Z vector containing the $5'$ and $3'$ *Xenopus* untranslated regions²¹.

cAMP and cGMP¹³. Finally, the maximum open probability at high concentrations of cGMP for the chimaera (0.85 ± 0.16 , $n=3$) is identical to the retinal channel (0.85 ± 0.12 , $n=5$) and distinct from the olfactory channel (0.42 ± 0.33 , $n=4$).

To determine whether H5 also controls ion selectivity through the cyclic nucleotide-gated channels, we measured the permeabil-

FIG. 3 Inorganic cation selectivity of the retinal, chimaeric and olfactory CNG channels. a, I - V relationships in response to ramp depolarizations for each channel. The patch pipette contained K^+ as its only cation; the bath solution contained a different metal test cation. Currents were normalized to the current at -20 mV. b, E_{rev} versus ion radius for the three channels. E_{rev} values were (mean $\pm$ s.d.) as follows. Retinal channel: Li^+ , 12.9 ± 1.2 ($n=5$); Na^+ , 2.8 ± 0.7 (5); K^+ , 0.0 ± 0.4 (5); Rb^+ , 10.6 ± 0.8 (5); Cs^+ , 27.1 ± 1.6 (5). Chimaeric channel: Li^+ , 3.7 ± 0.9 (5); Na^+ , 1.5 ± 1.0 (5); K^+ , 0.0 ± 0.8 (5); Rb^+ , 6.0 ± 0.6 (4); Cs^+ , 16.7 ± 1.4 (4). Olfactory channel: Li^+ , 1.7 ± 2.0 (7); Na^+ , 0.3 ± 0.7 (9); K^+ , 0.0 ± 1.0 (9); Rb^+ , 7.1 ± 0.5 (7); Cs^+ , 17.9 ± 1.6 (6). E_{rev} values are expressed relative to the K^+ reversal potential (set at 0 mV) to correct for metal-liquid junction potentials (<5 mV). All liquid-liquid junction potentials were calculated to be less than 0.5 mV.

METHODS. Currents recorded using a Heka EPC-9 patch-clamp amplifier, filtered at 1 kHz, digitized at 2 kHz and stored on an Atari computer. A 3 M KCl agar bridge was used as the ground electrode. Pipette contained (in mM): 125 K, 97 Cl $^-$, 10 HEPES, 10 EGTA, 1 EDTA (adjusted to pH 7.2 with KOH). The bath solution was identical to the pipette solution, except that the test ion was substituted for K^+ . Ramps depolarized the patch from -80 mV to $+80$ mV at 1 mV/ 3 ms. For each test



ion, four control ramps were applied in the absence of cGMP followed by four test ramps in the presence of 1 mM cGMP and four more control ramps after washout of cGMP. The control ramps preceding cGMP were averaged and subtracted from those following cGMP washout to verify leak-resistance stability. If leak resistance changed, the data were dropped from the analysis. Otherwise all control ramps were averaged and subtracted from the averaged test ramps. E_{rev} determined by linear regression of I - V curves around zero current.

ities of a series of monovalent metal cations from bi-ionic reversal potentials (E_{rev}) with potassium as the external cation and the test cation (X) on the inside of the membrane. The permeability ratio for a given ion X to that of K^+ (P_X/P_K) was then determined from the Goldman-Hodgkin-Katz equation: $P_X/P_K = \exp(-E_{rev}F/RT)$. As shown in Fig. 3, the olfactory and retinal channels select weakly among the metal cations. The retinal channel permeability sequence was $K^+ > Na^+ > Rb^+ > Li^+ > Cs^+$, with permeabilities relative to K^+ of 1.0, 0.90, 0.66, 0.60 and 0.35. The olfactory channel shows a slightly different sequence: $K^+ \geq Na^+ > Li^+ > Rb^+ > Cs^+$, with lower relative K^+ selectivity ratios of 1.0, 0.99, 0.94, 0.75, 0.50. The metal cation selectivity of the retinal-olfactory chimaeric channel was nearly identical to that of the olfactory channel (Fig. 3), with a permeability sequence of: $K^+ \geq Na^+ > Li^+ > Rb^+ > Cs^+$ and relative permeability ratios of 1.0, 0.94, 0.86, 0.79, 0.52.

The larger conductance and lower ionic selectivity of the olfactory channel relative to the retinal channel suggest that the olfactory channel pore may be wider than that of the retinal channel. To determine the apparent diameter we measured the relative permeabilities of the retinal and olfactory channels to a sequence of methylated ammonium cations of increasing diameter applied

to the inside of the membrane^{19,20}. As shown in Fig. 4, the reversal potential for dimethylammonium (diameter 4.6 Å) for the retinal channel (72.3 ± 5.3 mV, $n=5$) is significantly more positive than the reversal potential for the olfactory channel (57.9 ± 6.8 , $n=7$; $P < 0.005$, t -test), indicating a lower permeability. The larger cation trimethylammonium (diameter of 5.2 Å) fails to permeate the retinal channel even at voltages up to +120 mV (Fig. 4b).

In contrast, the olfactory channel is significantly permeable to trimethylammonium which carries a clear outward current (mean reversal potential of 97.4 ± 5.8 mV; $n=4$). We fitted plots of relative permeability versus organic cation diameter with an equation that relates ionic permeability to pore diameter based on a hydrodynamic model for ion movements¹⁹ (Fig. 4c). Although this model is an oversimplification of ion permeation, it does provide a reasonably good fit to the data, yielding estimates for apparent pore diameter of 5.8–5.9 Å for the retinal channel and 6.3–6.4 Å for the olfactory channel.

To investigate whether the H5 domain determines pore diameter, we measured the organic cation permeability of the retinal-olfactory chimaera. The data clearly indicate that the chimaeric channel has a large pore, similar to that of the olfac-

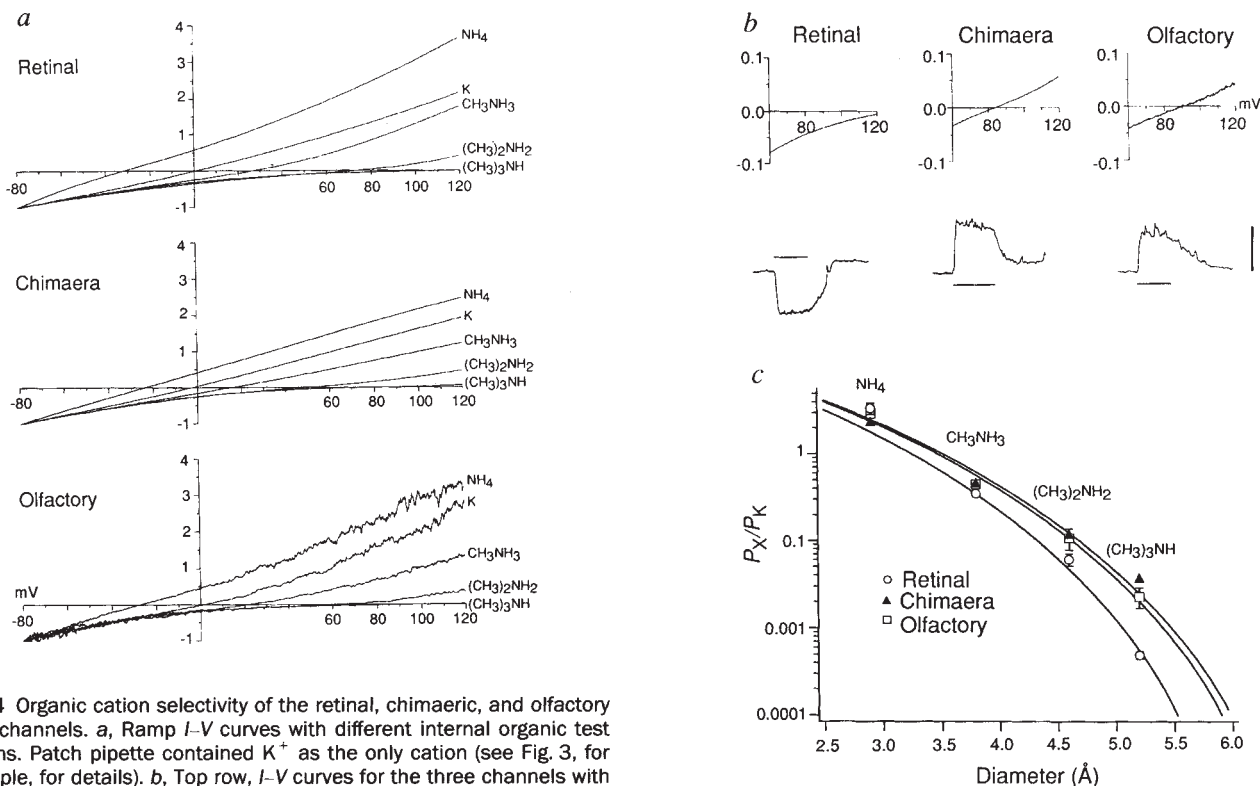


FIG. 4 Organic cation selectivity of the retinal, chimaeric, and olfactory CNG channels. **a**, Ramp I - V curves with different internal organic test cations. Patch pipette contained K^+ as the only cation (see Fig. 3, for example, for details). **b**, Top row, I - V curves for the three channels with trimethylammonium as the internal cation. Only a portion of the I - V curve is shown to demonstrate the position of E_{rev} . Bottom row, response to cGMP (1 mM; bars) with trimethylammonium as the internal cation at a constant potential of +100 mV. cGMP elicits inward current for the retinal channel and outward current for the olfactory and chimaeric channels. Scale bar, 10 pA for retinal and olfactory channel patches; 25 pA for chimaera. Bar, 1 s application for retinal and olfactory channel patches; 7 s application for chimaera. **c**, P_X/P_K versus cation radius for the three channels. The superimposed curves are best fits (χ^2) to the following hydrodynamic equation¹⁹: $P_X/P_K = C(1-a)^2(1-2.105a+2.0865a^3-1.7068a^5+0.72603a^6)/(1-0.75857a^5)$, where $a=q/d$, q is the diameter of cation X and d is the pore diameter. C is a constant that was either allowed to be different for the three fits (yielding pore diameters of 5.8, 6.4 and 6.8 Å, for retinal, olfactory and chimaeric channels, respectively) or fixed to a single value (yielding diameters of 5.9, 6.3 and 6.4 Å; fits shown with fixed C). Mean E_{rev} ($\pm$ s.d.) were as follows, retinal channel: NH_4^+ , -31.0 ± 4.1 (7); K^+ , 0.0 ± 2.5 (7);

$CH_3NH_3^+$, 26.8 ± 1.6 (7); $(CH_3)_2NH_2^+$, 72.3 ± 5.2 (5); $(CH_3)_3NH^+$, 136.6 ± 1.2 (4). Chimaeric channel: NH_4^+ , -22.2 ± 1.6 (6); K^+ , 0.0 ± 2.1 (6); $CH_3NH_3^+$, 19.1 ± 0.7 (6); $(CH_3)_2NH_2^+$, 54.4 ± 1.3 (5); $(CH_3)_3NH^+$, 84.5 ± 5.4 (4). Olfactory channel: NH_4^+ , -27.6 ± 1.6 (9); K^+ , 0.0 ± 0.7 (15); $CH_3NH_3^+$, 20.4 ± 1.6 (10); $(CH_3)_2NH_2^+$, 57.9 ± 6.8 (7); $(CH_3)_3NH^+$, 97.4 ± 5.8 (4). Liquid junction potentials were calculated to be less than 0.5 mV. For the retinal channel, E_{rev} for $(CH_3)_3NH^+$ was extrapolated from fits of Goldman-Hodgkin-Katz current equation²² to I - V curves between +80 mV and +120 mV. Ionic radii of organic cations were estimated from the smallest cylinder into which Corey-Pauling-Koltum models would fit^{19,20}. Radius of NH_4^+ assumes that the hydrogen atoms form hydrogen bonds. For calculations of relative channel K^+ conductances (see text), the diameter of partially hydrated K^+ (4.9 Å) is the effective diameter of an oblate ellipsoid with rectangular dimensions of a K^+ ion in contact with one water molecule (5.4 Å $\times$ 4.1 Å).

tory channel. Thus, the chimaeric channel is permeable to trimethylammonium which carries a clear outward current (mean reversal potential 84.5 ± 5.4 mV; $n=4$). Fits of the hydrodynamic equation yield an apparent pore diameter of $6.4\text{--}6.8$ Å, similar to the diameter of the olfactory channel.

Assuming that K^+ permeates the channel in contact with a single water molecule (effective diameter of 4.9 Å; see legend to Fig. 4), we used the hydrodynamic equation to predict how the K^+ conductance would vary among the retinal, olfactory and chimaeric channels. The relative slope conductances for the retinal : olfactory : chimaeric channels are $1.0 : 2.4 : 3.2$ and $1.0 : 3.3 : 3.7$, measured at $+80$ mV and -80 mV respectively. These ratios are in good agreement with the ratios predicted from the hydrodynamic equation of $1 : 2.7 : 3.3$.

These results indicate that ion permeation through the cyclic nucleotide-gated channels is largely determined by the H5 domain, similar to voltage-gated channels. Moreover, our results indicate that differences in ion permeation through retinal and olfactory cyclic nucleotide-gated channels are largely determined by differences in pore diameter, with specific ion binding playing a lesser role. In this respect, the cyclic nucleotide-gated channels are likely to differ from the more highly selective voltage-gated K^+ , Na^+ and Ca^{2+} channels, in which specific ion binding is probably more important. Finally, our results provide direct evidence that H5 determines pore diameter, strongly supporting

the view that this domain constitutes the selectivity filter of the pore. □

Received 8 March; accepted 15 April 1993.

1. Hille, B. *Ionic Channels of Excitable Membranes* (Sinauer, Sunderland, Massachusetts 1991).
2. MacKinnon, R. & Miller, C. *Science* **245**, 1382–1385 (1989).
3. MacKinnon, R. & Yellen, G. *Science* **250**, 276–279 (1990).
4. MacKinnon, R., Heginbotham, L. & Abramson, T. *Neuron* **5**, 767–771 (1990).
5. Yellen, G., Jurman, M. R., Abramson, T. & MacKinnon, R. *Science* **251**, 939–942 (1991).
6. Hartmann, H. A. et al. *Science* **251**, 942–944 (1991).
7. Yool, A. J. & Schwarz, T. L. *Nature* **349**, 700–704 (1991).
8. Kirsch, G. E. et al. *Neuron* **8**, 499–505 (1992).
9. Noda, M., Suzuki, H., Numa, S. & Stuhmer, W. *FEBS Lett.* **259**, 213–216 (1989).
10. Terlau, H. et al. *FEBS Lett.* **293**, 93–96 (1991).
11. Satin, J. et al. *Science* **256**, 1202–1205 (1992).
12. Heinemann, S. H., Terlau, H., Stuhmer, W., Imoto, K. & Numa, S. *Nature* **356**, 441–443 (1992).
13. Backx, P. H., Yue, D. T., Lawrence, J. H., Marban, E. & Tomaselli, G. F. *Science* **257**, 248–251 (1992).
14. Kaupp, B. U. et al. *Nature* **342**, 762–766 (1989).
15. Goulding, E. H. et al. *Neuron* **8**, 45–58 (1992).
16. Jan, L. Y. & Jan, Y. N. *Nature* **345**, 672 (1990).
17. Heginbotham, L., Abramson, T. & MacKinnon, R. *Science* **258**, 1152–1155 (1992).
18. Isacoff, E. Y., Jan, Y. N. & Jan, L. Y. *Nature* **353**, 86–90 (1991).
19. Dwyer, T. M., Adams, D. J. & Hille, B. *J. gen. Phys.* **75**, 469–492 (1980).
20. McCleskey, E. W. & Almers, W. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7149–7153 (1985).
21. Liman, E. R., Tytgat, J. & Hess, P. *Neuron* **9**, 861–871 (1992).
22. Hodgkin, A. L. & Katz, B. *J. Physiol.* **108**, 37–77 (1949).

ACKNOWLEDGEMENTS. We thank R. Axel for support and advice, B. Hille for discussion, R. Kramer and J. Ngai for comments on the manuscript, E. Liman for the expression vector, S. Mack and C. Lamb for help with artwork, J. Riley and L. Eoyang for technical assistance, and J.-p. Cong for DNA sequencing. G.R.T. thanks Sanity for initial support.

An early pre-liver intra-embryonic source of CFU-S in the developing mouse

Alexander L. Medvinsky*, Nina L. Samoylina*, Albrecht M. Müller† & Elaine A. Dzierzak†

* Laboratory of Physiology of Hematopoiesis, National Research Centre for Hematology, Novozykovsky pr. 4a, 125167, Moscow, Russia

† National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

It is widely accepted that during murine embryogenesis, totipotent haematopoietic stem cells first originate in the yolk sac, then migrate to the fetal liver and finally colonize the bone marrow shortly before birth^{1,2}. This view is based on *in vitro* studies showing that yolk sac cells can differentiate into various haematopoietic lineages^{1,3–7} and *in vivo* studies showing that yolk sac contains spleen colony-forming units (CFU-S) beginning at day 8 of gestation¹. However, some investigators have failed to find statistically significant numbers of CFU-S arising from day 9 yolk sac^{3,8–11} and, although one group reported that yolk sac could repopulate the haematopoietic system of W mutant mice², others have failed to confirm yolk sac-derived repopulation of adults^{3,12}. In the avian and amphibian systems, the yolk sac gives rise only to early, transitory haematopoiesis whereas the definitive adult haematopoietic stem cells in these vertebrates are derived from the mesodermal region containing the dorsal aorta^{13–17}. Because this analogous area of the mouse embryo has not been previously examined for haematopoietic activity, we directly compared the CFU-S activity of the aorta, gonad, mesonephros (AGM) region with the yolk sac and fetal liver during embryogenesis. Here we report that this intra-embryonic AGM region contains CFU-S activity at a higher frequency than that in embryonic yolk sac and that such activity appears in the AGM region before the fetal liver.

To determine the location and temporal appearance of CFU-S₈ activity in the mouse embryo, tissues were dissected from 14

somite pair (sp) stage embryos (day 8 p.c.) up to the 36–37 sp stage. As shown in Fig. 1a, significant numbers of CFU-S₈ were observed in the yolk sac and embryo body beginning at the 26–27 sp stage (day 9.5 p.c.). This stage-specific analysis indicates that through day 10 p.c., the numbers of CFU-S₈ rise more rapidly and are more abundant in the embryo body than in the yolk sac.

Specific microdissections of the embryo body were done to determine if CFU-S₈ activity on day 10 of development could be further localized. The highest concentration of cells contributing to CFU-S₈ were in the axial area of the embryo body comprising the AGM region (Fig. 1b). CFU-S were first detected in this region at the 31–33 sp stage, the earliest time point that this tissue can be accurately dissected from surrounding tissues. Their number continued to rise to a maximum at the 38–40 sp stage, and then fell rapidly by day 11. The frequency of CFU-S₈ in the AGM region on day 10 of gestation was as high as that in adult bone marrow, reaching $20\text{--}40$ per 10^5 cells (Fig. 1c). Lower concentrations of CFU-S₈ were found in the yolk sac until day 11 and were comparable to those previously reported¹. As we have shown^{10,11}, no CFU-S₈ were detected in the liver rudiment until the end of day 10 p.c. (38–40 sp) when the AGM region had peaked in CFU-S activity (Fig. 1b, c). The frequency of CFU-S₈ remaining in the body tissues after excision of the AGM region and the liver rudiment did not exceed 1.85 per 10^6 cells during day 10 of development.

Because CFU-S₈ are rather mature progenitors in the haematopoietic hierarchy^{18–20}, we also tested the same embryonic tissues for the presence of more primitive progenitors by the day 11 CFU-S assay. As shown in Table 1, no donor-derived CFU-S₁₁ were found in any embryo tissues from 0–29 sp stages. On day 10 of gestation (31–35 sp) statistically significant numbers of CFU-S₁₁ were found only in the AGM region. Although the frequency of AGM-derived CFU-S₁₁ was high (107.6 ± 45.1 in 10^6 cells) and found to be comparable to that for CFU-S₈, only low numbers if any of CFU-S₁₁ (0.56 ± 0.29 in 10^6 cells) were observed in the yolk sac at this stage of development. The total number of yolk sac-derived CFU-S₁₁ was 20 times less and the frequency 200 times less than found for the AGM region. Significant colony activity begins to appear in the yolk sac only by the end of day 10 p.c. No CFU-S₁₁ were found in the liver up

TABLE 1 CFU-S₁₁ activity in the tissues of day 8–11 mouse embryos

Age (days)	8*	9†	9‡	10§	10	11	11
Somite stage	0–17	17–24	24–29	31–35	38–40	41–47	>48
Yolk sac and body	1¶/23						
Whole body		2¶/44	1¶/46				
Yolk sac		0 /43	1¶/45	0.06±0.03	0.34±0.18	2.90±0.84	3.89±0.72
Liver		0 /29	0 /53	0	0.02±0.03	1.41±1.00	4.84±0.79
AGM region				1.2 ±0.4	0.94±0.42	5.98±0.87	2.25±1.06

Data for *, †, ‡ are represented as total number of CFU-S/number of embryos. Preparation of embryonic tissues (*, †, §) was as in Fig. 1 except that 10 embryos (*) and 16 embryos (†) were not treated with collagenase.

* Pooled results from two separate experiments. ¶ Denotes testing for the T6 donor chromosome marker. No evidence of donor T6 marked cells was found indicating this as an endogenous recipient derived colony.

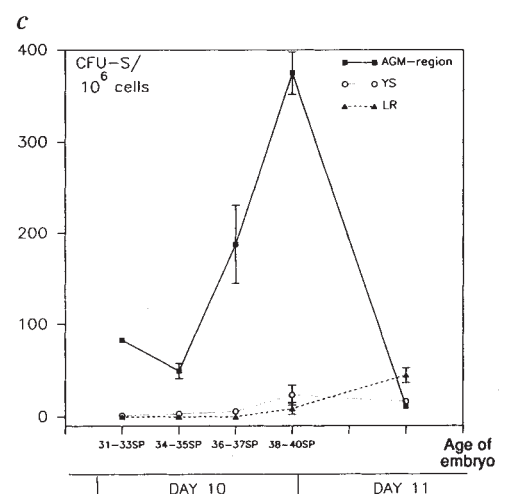
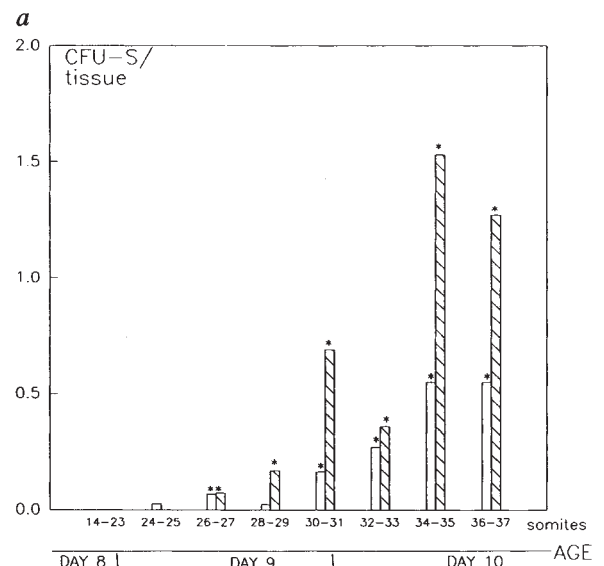
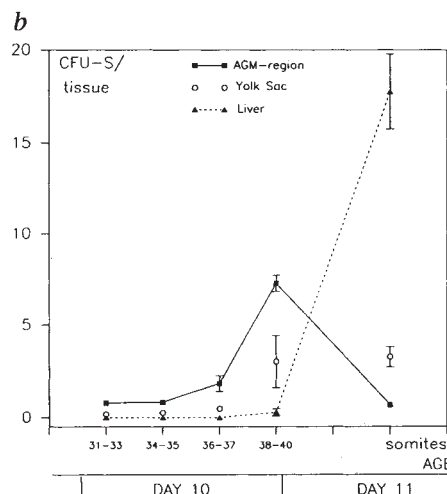
† Pooled results from five separate experiments. ¶ Denotes testing for the T6 donor chromosome marker as above. No donor cells were found to contribute to the two colonies.

‡ Embryonic tissues were prepared as in Fig. 1 except that the donor embryos were obtained from matings of homozygous globin transgenic males (ref. 32) with (CBA × C57BL/10)F₁ females. Cells were disaggregated in 0.1% collagenase with 20% FCS and injected into ⁶⁰Co irradiated (1,000 RADS, split dose) female (CBA × C57BL/10)F₁ mice. Pooled results from three separate experiments are shown. ¶ Denotes testing for the globin transgenic marker. No donor cells were found to contribute.

§ Data for day 10 and 11 embryos shows the number of CFU-S per tissue. A total of 43 (31–35 sp), 26 (38–40 sp), 26 (41–47 sp) and 10 (>48 sp) embryos were used. The means and standard errors of two or three experiments are presented. The difference between body remnants and control groups is insignificant.

FIG. 1 CFU-S₈ activity in yolk sac and embryonic tissues. *a*, histogram analysis of total CFU-S₈ in the yolk sac (open bars) and body (striped bars) of the developing mouse embryo. Bars represent weighted mean values. An asterisk indicates that the value is significantly different ($P < 0.05$) from the irradiated control by Smirnov's homogeneity criterion for two samples³¹. Three to eight experiments were done for each experimental point. The total number of injected yolk sacs (*n*) and bodies (*m*) at each experimental point are: 14–23, *n*=32, *m*=32; 24–25, *n*=75, *m*=59; 26–27, *n*=72, *m*=78; 28–29, *n*=66, *m*=74; 30–31, *n*=41, *m*=25; 32–33, *n*=32, *m*=32; 34–35, *n*=46, *m*=34; 36–37, *n*=40, *m*=40. No more than 1 colony per 55 spleens was found in irradiated control groups of mice. *b*, Number of CFU-S₈ per embryo and *c*, CFU-S₈ frequency in 10⁶ cells in the AGM region, yolk sac and liver at days 10–11 p.c. Error bars indicate standard error of the mean (s.e.m.). Where no error bars are visible s.e.m. falls within the symbol. Two to seven experiments were done for each experimental point. The numbers of injected embryos at each point are: 31–33, *n*=18; 34–35, *n*=57; 36–37, *n*=42; 38–40, *n*=17; day 11, *n*=54.

METHODS. 8–10-day C57BL/6 × CBA embryos (the day of discovery of vaginal plug was counted as day 0) were removed from uteri and dissected with fine needles in L-15 medium (Flow Lab.) with 5% fetal bovine serum (Flow Lab.) under the dissecting microscope. For *a*, yolk sacs and embryo bodies, after the removal of liver rudiments (after 16 sp stage), were cut into large fragments and incubated with 0.04% collagenase (Sigma cat. no. C9407) in L-15 medium with 5% of fetal bovine serum. *b*, *c*, AGM regions, yolk sacs and liver rudiments were also extracted with fine needles and collagenase treated. Collagenase digestion permits the enhanced recovery of early haematopoietic precursors from mouse yolk sac⁵ and was most probably responsible for our observations of the earlier appearance of CFU-S in this study than we have previously reported^{10,11}. After gentle pipetting and washing, cells were injected into irradiated mice for CFU-S assay. Recipients were exposed to 13.6 Gy of total irradiation separated by 3-h intervals at a dose rate of 0.176 Gy per min from a ¹³⁷Cs source. Eight days after injection, recipients were killed, spleens were fixed in Bouin's solution and macroscopic colonies were counted.



to the 38–40 sp stage and only insignificant numbers were observed in the body remnants after tissue dissection. To determine whether CFU-S₁₁ from day 9 to 11 embryos were donor derived, a human globin transgenic marker was used. The transgenic marker was detected only after day 10 p.c. in AGM- and day 11 p.c. in AGM-, yolk sac- and liver-derived CFU-S₁₁ (Fig. 2). No donor signal was found in colonies derived from whole body or yolk sac cells at the 24–29 sp stage (see Table 1) indicating that they were host derived.

The dramatic rise in CFU-S content in the liver rudiment (day 11 p.c.) coincides with a drop in the number of AGM region-derived CFU-S (Fig. 1 and Table 1) and corresponds with the slightly earlier presence of colony activity in the circulation of 37–38 sp embryos¹⁰. Before the presence of CFU-S activity in the liver rudiment^{11,21,22}, erythropoietic and committed progenitor cells have been detected. Thus, the liver rudiment appears to be colonized by two successive waves of haematopoietic cells; the first from the yolk sac¹ which is rich in committed precursors and the second wave by immature stem cells from the AGM region and yolk sac.

Histological examination of a day 10 p.c. embryo was also done to determine whether the AGM region contained haematopoietic cells (Fig. 3). Despite the presence of CFU-S activity and primordial germ cells, no haematopoietic foci were found. Note that stem cell factor (SCF), which plays an important role in controlling germ-cell development, haematopoiesis and CFU-S production^{23–26}, is highly expressed in the AGM region at days 9–10 p.c.^{27,28} whereas yolk sac expression is low²⁷. This pattern of SCF expression correlates well with the levels of CFU-S activity we have found in the AGM region and the yolk sac. Thus SCF may be responsible for temporal and spatial patterns of CFU-S distribution in the embryo, although it is unclear whether

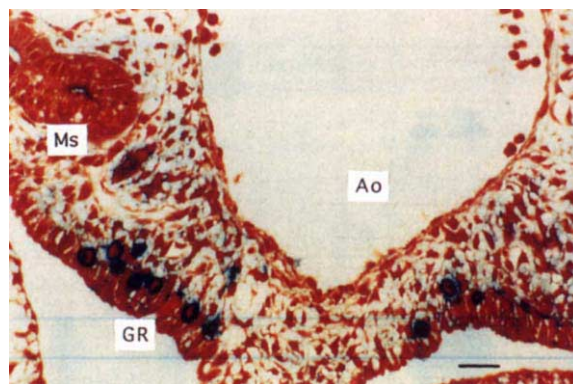


FIG. 3 Tissue section of the AGM region of a 34 sp mouse embryo (10 days of gestation). A transverse section of the AGM region was embedded in plastic for visualization of haematopoietic foci and stained for alkaline phosphatase activity. Germ cells bearing alkaline phosphatase activity appear with blue cytoplasm. At this stage the AGM region harbours both early haematopoietic and germ cell progenitors. Ms, mesonephros; GR, genital ridge; Ao, dorsal aorta. Scale bar, 6 μ m. METHODS. Embryos were fixed at 4 °C in 4% formaldehyde, PBS (pH 7.4). After overnight washing in cold PBS (three changes) specimens were embedded in Histo-resin (Leica) at 4 °C. The blocks were sectioned with a steel knife with tungsten carbide edge (Leica) on a Reichert-Young 2030 microtome at 2 μ m. Sections were stained for 1.5 h at 37 °C for alkaline phosphatase activity with Sigma kit No. 86-C and counterstained with neutral red.

CFU-S mature from their progenitors *in situ* or migrate²⁹ to the AGM from some external source.

We have shown here that the AGM region contains a higher frequency of CFU-S activity than the yolk sac. Although the CFU-S assay demonstrates the presence of haematopoietic progenitors in the AGM region and yolk sac, it does not specify either of these tissues as a primary source of pluripotential haematopoietic stem cells. Previously, using an assay to detect pre-CFU-S⁸, we have reported activity in both yolk sac and embryo body at the 24–25 sp stage¹¹. Therefore, to test for the most primitive stem cells^{20,30}, we are presently comparing the long-term repopulating activity of transgene-marked embryonic tissues. Thus, before the onset of definitive haematopoiesis in embryonic liver, there exists in parallel with the yolk sac another source of CFU-S activity within the embryo body. The AGM region is an abundant source of haematopoietic activity which rapidly surpasses that in the yolk sac at early stages and may be analogous to the intra-embryonic site of definitive haematopoiesis in avian and amphibian embryos. □

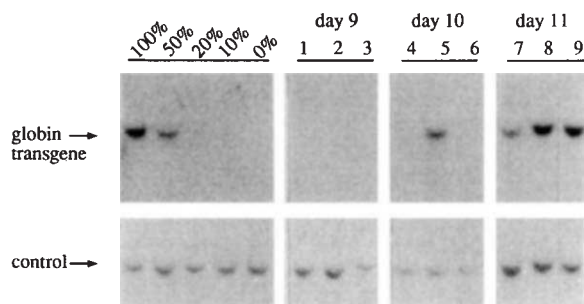


FIG. 2 Southern blot of CFU-S₁₁ for determination of donor contribution. CFU-S₁₁ were dissected from recipient animals which received an injection of day 9 (24–29 sp), 10 (35–39 sp) or 11 (greater than 40 sp) marked embryonic tissues: lanes 1, 4, 8 from yolk sac, lane 2 from embryo body, lanes 5, 7 from AGM, lanes 3, 9 from liver and lane 6 from body remnants. Human globin transgenic marked embryos were derived from homozygous transgenic (line 72) males³² mated with normal (CBA × C57BL/10) females. Embryonic cells were prepared and transplanted as described in Table 1. Individual spleen colonies were dissected and DNA prepared at day 11 post-transplantation as shown in lanes 1, 2 and 4–9. Whole spleen DNA was prepared for lane 3 because no colonies arose from injected day 9 liver cells. DNA was cut with EcoRI (8 μ g) and blotted as previously described³² and a 3.3 kilobase (kb) EcoRI fragment from the 5' end of the human beta-globin locus (LCR) was nick-translated and used as a probe. Control DNAs were a mixture (percentage as indicated) of DNA obtained from a heterozygous transgenic mouse and a non-transgenic mouse spleen. Donor contribution was quantified by densitometric analysis of autoradiograms (Chromoscan 3). The single-copy endogenous Thy-1 gene served as a loading control after reprobing with a 1.3 kb Apal Thy-1 gene fragment³³. Signals in lane 5, 8 and 9 show that these CFU-S are 100% donor derived. The DNA in lane 7 is only 50% donor derived and may be the result of imprecise dissection of this spleen colony. Longer exposures show no signal in any other lanes.

Received 19 April; accepted 27 May 1993.

- Moore, M. A. S. & Metcalf, D. *Br. J. Haematol.* **18**, 279–296 (1970).
- Toles, J. F. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 7456–7459 (1989).
- Sonoda, T. *et al. Devl Biol.* **97**, 89–94 (1983).
- Johnson, G. R. & Barker, D. C. *Exptl Haematol.* **13**, 200–208 (1985).
- Wong, P. M. C. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 3851–3854 (1986).
- Eren, T. *et al. Immun. Res.* **6**, 279–287 (1987).
- Liu, C. P. & Auerbach, R. *Development* **113**, 1315–1323 (1991).
- Perah, G. & Feldman, M. J. *Cell Physiol.* **91**, 193–200 (1977).
- Symann, M. *et al. Exptl Haematol.* **6**, 749–759 (1978).
- Samoylina, N. L. *et al. Sov. J. dev. Biol.* **21**, 127–133 (1990) (Engl. transl.).
- Samoylina, N. L., Medvinsky, A. L., Gan, O. I. & Manakova, T. E. in *Haematology and Blood Transfusion 35, Modern Trends in Human Leukemia IX* (eds Neth, R. *et al.*) 106–109 Springer, Berlin, New York, 1992).
- Harrison, D. E. *et al. Blood* **54**, 1152–1157 (1979).
- Dieterlen-Lievre, F. J. *Embryol. exp. Morph.* **33**, 607–619 (1975).
- Dieterlen-Lievre, F. & Martin, C. *Devl Biol.* **88**, 180–191 (1981).
- Turpen, J. B. *et al. Devl Biol.* **85**, 99–112 (1981).
- Maeno, M. *et al. Devl Biol.* **110**, 503–508 (1985).
- Cormier, F. & Dieterlen-Lievre, F. *Development* **102**, 279–285 (1988).
- Magli, M. C. *et al. Nature* **295**, 527–529 (1982).
- Wolf, N. C. & Priestley, G. V. *Exptl Haematol.* **14**, 676–682 (1986).
- Ploemacher, R. E. & Brons, N. H. C. *Exptl Haematol.* **16**, 21–26 (1988).
- Houssaint, E. *et al. Cell Diff.* **10**, 243–252 (1981).

22. Johnson, G. R. & Jones, R. O. *J. Embryol. exp. Morph.* **30**, 83–96 (1973).
 23. Witte, O. N. *Cell* **63**, 5–6 (1990).
 24. Russell, E. S. *Adv. Genet.* **20**, 357–459 (1979).
 25. De Vries, P. et al. *J. exp. Med.* **173**, 1205–1211 (1991).
 26. Bodine, D. M. et al. *Blood* **79**, 913–919 (1992).
 27. Matsui, Y. et al. *Nature* **347**, 667–669 (1990).
 28. Keshet, E. et al. *EMBO J.* **10**, 2425–2435 (1991).
 29. Blume-Jensen, P. et al. *EMBO J.* **10**, 4121–4128 (1991).
 30. Jones, R. J. et al. *Nature* **347**, 188–189 (1990).

31. Bolshov, L. N. & Smirnov, N. V. in *Tables of Mathematical Statistics* 464 (Nauka, Moscow, 1965).
 32. Strouboulis, J. et al. *Genes Dev.* **6**, 1857–1864 (1992).
 33. Giguere, V. et al. *EMBO J.* **8**, 2017–2024 (1983).

ACKNOWLEDGEMENTS. We thank J. Strouboulis and F. Grosfeld for the globin transgenic mice and J. Cherkov, J. Dick and F. Grosfeld for helpful discussions and critically reviewing this manuscript. This work was supported by the Biotechnology Scientific Council of the Russian Academy of Medical Sciences and the MRC, UK.

Para-aortic splanchnopleura from early mouse embryos contains B1a cell progenitors

Isabelle E. Godin*, Juan A. Garcia-Porrero†, Antonio Coutinho‡, Françoise Dieterlen-Lièvre* & Miguel A. R. Marcos‡§

* Institut d'Embryologie Cellulaire et Moléculaire du CNRS et du Collège de France, 94130 Nogent-sur-Marne, France

† Department of Anatomy and cell Biology, Universidad de Cantabria, Santander 39011, Spain

‡ Unité d'Immunobiologie CNRS UA 357, Institut Pasteur, Paris 75015, France

§ Department of Medicine, Universidad de Salamanca, Salamanca 37007, Spain

DEFINITIVE erythropoiesis in birds originates from stem cells that emerge in the splanchnopleural mesoderm near the embryonic aorta^{1–4}. The yolk sac is still generally held to be the unique provider of haematopoietic stem cells during mammalian ontogeny⁵, although there may be an alternative intraembryonic source of stem cells in the mouse fetus^{6,7}. Here we search for a possible non-yolk-sac source of stem cells by grafting intraembryonic splanchnopleura from 10- to 18-somite mouse embryos into adult immunodeficient SCID mice. We find significant amounts of donor-derived serum IgM, normal numbers of IgM-secreting plasma cells, and the B1a (IgM^{bright}B220^{dull}CD5⁺) cell subset to be fully reconstituted by donor progenitors 3 to 6 months after engraftment. The haematogenic capacity revealed in our experiments is present in a previously unrecognized site, the earliest described in the embryo, 12 hours before fetal liver colonization.

The splanchnopleura region was dissected from day 8.5–9 BALB/c (H-2^d, Igh-6a) embryos (10–18 somites; Fig. 1A, B), and grafted under the kidney capsule of CB.17-SCID (H-2^d, Igh-6b) mice. After 2–4 months of follow-up, significant amounts of donor-derived IgM^a (Fig. 2) were detected in mice (6 out of 15) grafted with embryonic splanchnopleura. No recipient of yolk sac ($n=10$), or younger embryos (<5 somites, $n=8$) showed seric IgM reconstitution in our experimental conditions. Some degree of 'leaky' endogenous Igh-6b was observed in the serum of 4–7-month-old SCID controls, as frequently happens in our colony; this endogenous contribution was reduced in the Igh-6a-reconstituted mice (Fig. 2). The mice were killed and it was found that the splanchnopleura-grafted, donor Igh-6a-reconstituted mice had large, soft, lobulated masses at the graft site under the kidney capsule (Fig. 1C). Only fibrotic scars were detected in mice from the splanchnopleura-grafted group which were negative for IgM^a allotype and in mice from the two other experimental groups. Counting cell numbers from lymphoid organs and body cavities did not reveal major changes in splanchnopleura-reconstituted versus control SCID mice, except for an almost normal mononuclear cell population present in the coelomic cavities (peritoneum, pleuropericardium) (Table 1a). Peyer's patches failed to develop in B-cell-reconstituted mice, and neither size nor cellularity of the rudimentary thymuses was increased.

Figure 3 shows a representative fluorescence-activated cell

sorter (FACS) profile of the lymphoid cell reconstitution in splanchnopleura-grafted hosts. No donor-derived surface IgM^{a+} B cells were detected in those grafted recipients (splanchnopleura, yolk sac, <5 somite embryos) that were negative for the same allotype in plasma. The peritoneal (and pleuropericardial; data not shown) exudates of the reconstituted mice revealed a homogeneous population displaying the surface antigen pattern characteristic of B1a lymphocytes⁸ (Fig. 3, two top rows). Although the cell complexity of these lymphocytes (as measured by side scatter) is comparable to that of equivalent cells from normal mice, their mean size (forward scatter) is larger (Fig. 3, contour plots). Coelomic exudates were basically devoid of CD5⁺ T cells, as well as of B1b and conventional B2 lymphocytes. In the spleen, the reconstituted mice contained small populations of IgM^{a+} lymphocytes (1–5% of total splenocytes; Fig. 3, third row); this was also the case in tiny lymph nodes that were recovered from the same mice (data not shown). A few T cells were recovered from spleen (Fig. 3; third row), but we could not establish their donor-recipient origin in the experimental combination studied here and they could have been contributed by endogenous leakiness of the SCID defect. The use of other mouse donor-recipient combinations will allow us to elucidate T-cell as well as myeloid potentialities of para-aortic splanchnopleura.

Although absolute numbers of mature B cells were reduced, splenic IgM^a-secreting plasma cells were well reconstituted (as detected in IgM allotype-specific enzyme-linked immunosorbent (ELISA), spot assays)⁹, and were close to levels normal for congenic BALB/c controls (Table 1b). The skewed plasma cell/B lymphocyte ratio of B1a reconstituted mice confirms the

TABLE 1 Lymphoid cell reconstitution of splanchnopleura-grafted SCID mice and lymphoid cell numbers in controls

a, Absolute numbers ($\times 10^6$) of mononuclear cells in lymphoid organs and coelomic cavities

	Control SCID	Grafted SCID	BALB/c
Bone marrow	21.5 $\pm$ 7	18.3 $\pm$ 0.3	22.5 $\pm$ 3
Spleen	37 $\pm$ 15	14 $\pm$ 4.2	89.5 $\pm$ 2
Peritoneum	0.3 $\pm$ 0.4	4 $\pm$ 0.4	5.9 $\pm$ 3
Pleuropericardium	<0.1	0.6 $\pm$ 0.3	1.2 $\pm$ 0.8

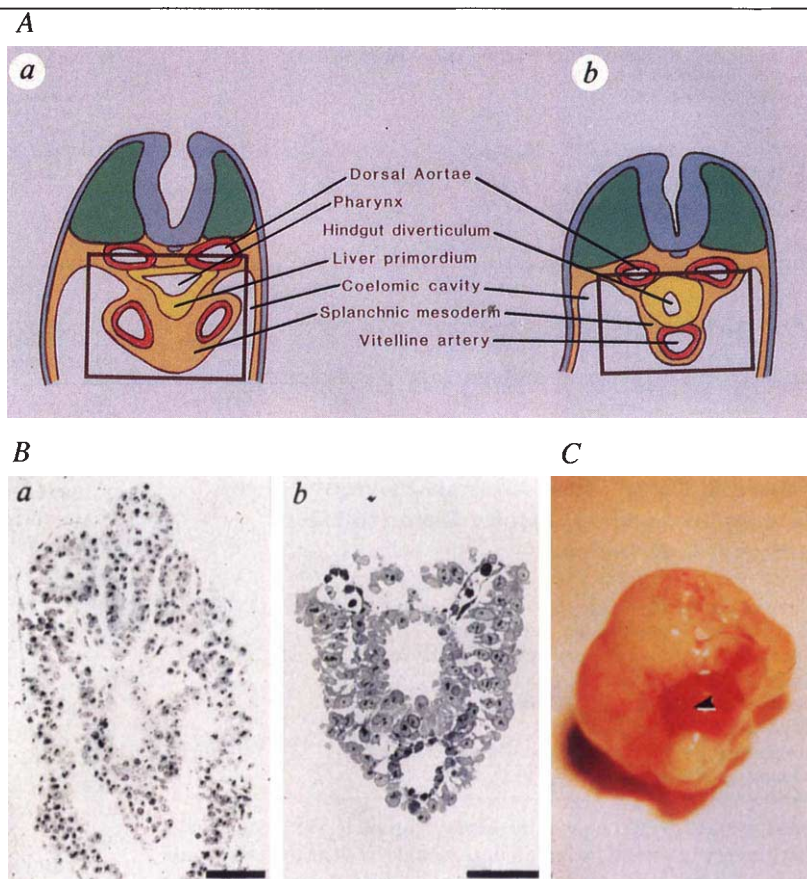
b, Donor-derived, surface IgM^{a+}, B lymphocytes and plasma cells in the spleen

	Surface IgM ^{a+} ($\times 10^6$)	IgM ^{a+} secretors ($\times 10^4$)	Plasma/B cell ratio ($\times 10^{-2}$)
BALB/c	35.2 $\pm$ 0.53	33.4 $\pm$ 9.8	0.95
Grafted SCID	0.79 $\pm$ 0.27	20 $\pm$ 16.3	20.7

a, Cell suspensions were prepared from solid organs by physical disruption and dead cells were excluded using trypan blue staining. Cell numbers correspond to two femurs per mouse (mean $\pm$ s.d.; $n=5-10$). b, The absolute numbers of IgM^{a+} B cells were obtained by referring the relative numbers of surface IgM^a cells, as obtained in flow cytometry windows that were negative in ungrafted SCID controls, to total splenocytes in each mouse analysed. IgM^a-secreting plasma cells were evaluated in allotype-specific ELISA spot assays^{9,18} (mean $\pm$ s.d.).

FIG. 1 A Schematic sections of a 10–18-somite embryo (dissected area is framed) at *a*, truncal level, comprising presumptive septum transversum, and *b*, abdominal level; B, semithin sections at abdominal level: *a*, before (scale bar, 100 μ m) and *b*, after (scale bar, 50 μ m) dissection of para-aortic splanchnopleura (mesoderm plus endoderm); C, outgrowth (yellowish) around the host kidney (red, arrow-head) 4 to 6 months after grafting in a SCID mouse.

METHODS. The splanchnopleura region was dissected from 8.5–9-day BALB/c (H-2^d, Igh-6a) embryos, and grafted under the kidney capsule of CB.17-SCID (H-2^d, Igh-6b) mice. Recipient mice were either untreated or irradiated with low-dose X-radiation (300 cGy total-body irradiation), without any modification in lymphoid reconstitution. Splanchnopleura from 2–8 developmentally matched embryos were introduced per mouse, each graft harbouring a mean of 1,000–3,000 cells. Age-matched fragments of YS, and whole earlier (<5 somites) YS-deprived embryos were also implanted for comparison. The developmental stage of donor embryos was defined by gestational age (day 0 represents day of post-coital plug detection) and number of somites.



preferential commitment of B1 cells to natural plasma cell differentiation previously suggested^{10–12}. Concerning the differentiation environment of donor-derived B cells in the SCID recipient, the bone marrow of those mice did not display any B-lineage change compared to ungrafted SCID, and no IgM^{a+} mature B cells could be found there (Fig. 3, bottom row). Virtually no bone marrow IgM^a-secreting cells were detected by analysis in ELISA spot assay (data not shown). Finally, the bone marrow of splanchnopleura-reconstituted SCID mice not only lacked B cells but also B-cell precursors derived from the

donor graft, as adoptive transfer of their bone marrow cells to 350 cGy-irradiated SCID mice did not generate any B lymphocytes (10⁷ total bone marrow cells i.v./mouse; analyses of recipients were undertaken 2–4 months after transfer; *n* = 5). We think that intermediate stages in B1a cell generation probably occurred *in situ* in the grafted tissues. We are currently analysing the renal subcapsular outgrowths for the presence of B-cell progenitors and/or the expression of pre-B-cell-specific genes.

Our results demonstrate the existence of B-cell progenitors in the mouse para-aortic splanchnopleura at the developmental stage of 10–18 somites (day 8.5–9 of gestation), a site identified as producing haematopoietic stem cells in the avian model^{1–4}. The mouse embryonic progenitors have an apparently restricted B lymphoid potential. Splanchnopleura grafted without dissociation gave rise to B1a lymphocytes only, a population which, although paramount in perinatal stages, is relatively reduced in the adult and restricted to the coelomic cavities¹³. Although we took extreme care to eliminate all blood cells, the grafts were obtained from embryos shortly after completion of vascular circulation which occurs at the 7-somite stage (F.D.-L., unpublished results) and they could eventually be contaminated by progenitors from yolk sac blood islands. Negative results were obtained with implants from this area under the same experimental conditions, in contrast with earlier demonstrations of the capacity of mouse yolk sac to generate mature B lymphocytes *in vivo* after their cell transfer i.v.¹⁴. Taking into account that our yolk-sac grafts contained 60–120 × 10³ cells (1–2 yolk sacs/SCID mouse), a very low number of progenitors or a failure of the yolk sac to develop under the kidney capsule might explain this discrepancy.

Most previous analyses used transfer of cell suspensions, whereas in our grafting protocol the implant retained its tissue organization. This procedure permits cellular interactions necessary for the expression and/or restriction of ultimate cell fates of undifferentiated progenitors. Why splanchnopleura had an apparently restricted lymphoid potential is still unclear. It is interesting to note that, despite the ectopic location of the graft,

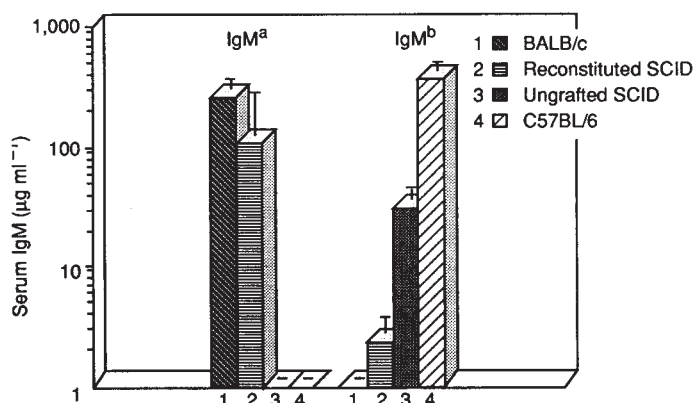


FIG. 2 Serum IgM levels (mean $\pm$ s.d.) in splanchnopleura-reconstituted SCID mice. Data are from ELISA studies undertaken in 96-well plastic plates coated with purified anti-IgM^a (RS3.1) (left columns: donor-derived reconstitution, and normal controls)¹⁹ or anti-IgM^b (MB.86) (right columns: endogenous SCID leakiness, and normal controls)²⁰ monoclonal antibodies¹⁸. The recipient SCID mice were checked at regular intervals by bleeding, and fully analysed between 2 and 5 months after grafting.

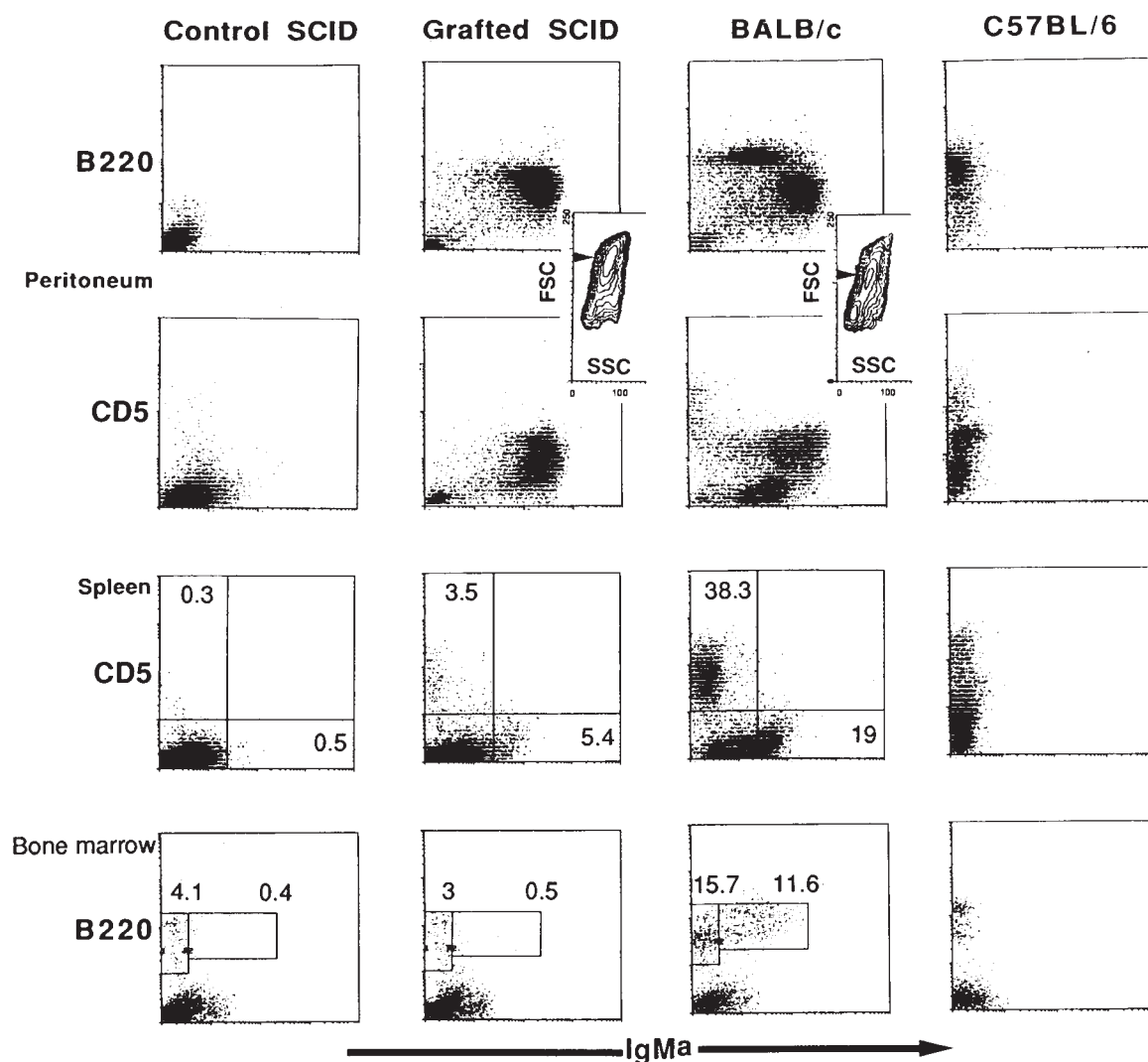


FIG. 3 Reconstitution of a homogeneous population of donor-type B1a lymphocytes in SCID mice grafted with embryonic para-aortic splanchnopleura. Single-cell suspensions prepared from peritoneal exudates, spleen and bone marrow were stained with fluorescein-(FITC-) conjugated anti-IgH-6a (RS3.1), and biotin-conjugated anti-B220 (RA3.3A1)²¹ or anti-CD5 (53-7.3)²² monoclonal antibodies. After a second incubation with streptavidin-phycoerythrin (Southern Biotechnology), a minimum of 10^5 cells per sample

were analysed in a Facscan flow cytometer using the LYSIS II program (Becton Dickinson). Isotype-matched irrelevant antibodies were used to define background fluorescence, and propidium iodide staining excluded dead cells. Data are displayed in dot plots (FL1 and FL2 fluorescence), where the intensity of fluorescence is shown in arbitrary units on a logarithmic scale, and in contour plots (FSC (forward scatter) and SSC (side scatter) light parameters; arrowheads point to the mean-sized peak of the population).

the reconstituted B1a cell population homed to its characteristic niche in the host. Restriction of the progeny from splanchnopleural progenitors to the B1a phenotype is striking, when compared with the derivatives from other tissues or cells. Day-13 omentum grafted under the kidney capsule gave rise to both B1a + B1b cells¹⁵, whereas precursors from fetal liver generated the whole B-cell compartment (B1a + B1b + B2)^{16,17}. Throughout the lifespan of the mouse, bone marrow progressively loses the ability to produce early B-cell populations¹⁵⁻¹⁷. These data support the existence of a developmentally established sequence of B-cell progenitor potentialities (B1a → B1b → B2). The findings reported here further suggest that the earliest intraembryonic haematopoietic progenitors, probably emerging and initiating their evolution *in situ* in the para-aortic splanchnopleura, become committed to the B1a lymphocyte generation. Whether this lymphoid restriction is imposed by a developmental program intrinsic to the haematopoietic stem cell, with the consequence that distinct stem cells would be required at different

times and/or locations, or whether haematopoietic stem cell fates are selected through cellular regulative interactions, is open to investigation. □

Received 30 November 1992; accept 16 March 1993.

- Dieterlen-Lièvre, F. *J. Embryol. exp. Morphol.* **33**, 607-619 (1975).
- Lassila, O., Eskola, J., Toivanen, P., Martin, C. & Dieterlen-Lièvre, F. *Nature* **272**, 353-354 (1978).
- Dieterlen-Lièvre, F. *Devl comp. Immun. suppl.* **3**, 75-80 (1984).
- Cormier, F. & Dieterlen-Lièvre, F. *Development* **102**, 272-285 (1988).
- Moore, M. A. S. & Metcalf, D. *Brit. J. Haemat.* **18**, 279-296 (1970).
- Marcos, M. A. R., Gutierrez, J. C., Huetz, F., Martinez-A., C. & Dieterlen-Lièvre, F. *Scand. J. Immun.* **34**, 129-135 (1991).
- Ogawa, M., Nishikawa, S., Yamamura, F., Naito, M. & Takahashi, K. *EMBO J.* **7**, 1337-1343 (1988).
- Kantor, A. B. *Immun. Today* **12**, 388-391 (1991).
- Sedgwick, J. & Holt, P. *J. Immun. Meth.* **57**, 301-303 (1983).
- Foster, I. & Rajewsky, K. *Eur. J. Immun.* **17**, 521-528 (1987).
- Kroese, F. G. M. et al. *Int. Immun.* **1**, 75-84 (1989).
- Marcos, M. A. R. et al. *Ann. N. Y. Acad. Sci.* **651**, 433-442 (1992).
- Hayakawa, K. & Hardy, R. R. *A. Rev. Immun.* **6**, 197-218 (1988).
- Paige, C. J., Kincade, P. W., Moore, M. A. S. & Lee, G. *J. exp. Med.* **150**, 548-563 (1979).
- Solvason, N., Luehen, A. & Kearney, J. F. *Int. Immun.* **3**, 543-550 (1991).

16. Hardy, R. R. & Hayakawa, K. *Proc. natn. Acad. Sci. U.S.A.* **88**, 11550-11554 (1991).
17. Kantor, A. B., Stall, A. M., Adams, S., Herzenberg, L. A. & Herzenberg, L. A. *Proc. natn. Acad. Sci. U.S.A.* **89**, 3320-3324 (1992).
18. Grandien, A. *et al. Int. Immun.* **4**, 1153-1160 (1992).
19. Schüppel, R., Wilke, J. & Weiler, E. *Eur. J. Immun.* **17**, 739-741 (1987).
20. Nishikawa, S., Yoshiyuki, S., Kina, T., Amagai, T. & Katsura, Y. *Immunogenetics* **23**, 137-139 (1986).
21. Coffman, R. L. & Weissman, I. L. *J. exp. Med.* **153**, 269-279 (1981).
22. Ledbetter, J. & Herzenberg, L. A. *Immunol. Rev.* **47**, 63-75 (1979).

ACKNOWLEDGEMENTS. We thank A. Cumano and M. L. Gaspar for discussion and suggestions, and E. Malenchère and J. Palacin for technical assistance. This work was supported by CNRS and by grants from FIS and DGICYT, and a fellowship from Société de Secours des Amis des Sciences to I.E.G.

Regulation of NMDA receptor phosphorylation by alternative splicing of the C-terminal domain

Whittemore G. Tingley, Katherine W. Roche, Angela K. Thompson & Richard L. Huganir*

Department of Neuroscience, Howard Hughes Medical Institute, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA

THE NMDA (*N*-methyl *D*-aspartate) receptors in the brain play a critical role in synaptic plasticity, synaptogenesis and excitotoxicity¹⁻³. Molecular cloning has demonstrated that NMDA receptors consist of several homologous subunits (NMDAR1, 2A-2D)⁴⁻⁷. A variety of studies have suggested that protein phosphorylation of NMDA receptors may regulate their function⁷⁻¹² and play a role in many forms of synaptic plasticity such as long-term potentiation^{13,14}. We have examined the phosphorylation of the NMDA receptor subunit NMDAR1 (NR1) by protein kinase C (PKC) in cells transiently expressing recombinant NR1 and in primary cultures of cortical neurons. PKC phosphorylation occurs on several distinct sites on the NR1 subunit. Most of these sites are contained within a single alternatively spliced exon in the C-terminal domain, which has previously been proposed to be on the extracellular side of the membrane^{4,5,15}. These results demonstrate that alternative splicing of the NR1 messenger RNA regulates its phosphorylation by PKC, and that mRNA splicing is a novel mechanism for regulating the sensitivity of glutamate receptors to protein phosphorylation. These results also provide evidence that the C-terminal domain of the NR1 protein is located intracellularly, suggesting that the proposed transmembrane topology model for glutamate receptors may be incorrect.

The phosphorylation of the NR1 subunit was examined directly by transiently transfecting human embryonic kidney 293 (HEK) cells with the complementary DNA for the NR1 subunit. These cells expressed the NR1 subunit, which could be detected as a 120K protein by immunoblotting cell extracts with antibodies directed against its C terminus (Fig. 1a). This 120K protein was also recognized by an N-terminal antibody (data not shown). To examine the phosphorylation of the NMDA receptor, cells were prelabelled with [³²P]orthophosphate, and the NR1 protein was immunoprecipitated with a combination of the N- and C-terminal antibodies. In the absence of activators of protein kinases, the basal phosphorylation of the NR1 subunit was significant; however, upon the activation of PKC with phorbol esters, the phosphorylation of the NR1 subunit was increased (Fig. 1b). Phosphopeptide maps of tryptic digests of the phosphorylated NR1 subunit using two-dimensional thin layer chromatography (TLC) demonstrated that PKC phosphorylation occurred on eight major tryptic phosphopeptides (Fig. 2a). Phorbol esters increased the phosphorylation of all of these sites

(data not shown). This phosphorylation occurred mostly on serine residues with a trace amount of phosphorylation occurring on threonine residues (Fig. 3c).

To determine whether the NR1 subunit is phosphorylated by PKC in neurons, we examined primary cultures of rat cortical neurons. Immunoblots of membranes isolated from these cultures or from adult rat forebrain using NR1 antibodies detected proteins migrating at about 120K (Fig. 1c). High-resolution immunoblots revealed that NR1 migrated as a triplet of about 120K proteins (data not shown). It is unclear whether these proteins represent splice variants, post-translational modification or partial proteolysis of the NR1 subunit. When neuronal cell cultures were prelabelled with [³²P]orthophosphate, the NR1 antibodies specifically immunoprecipitated phosphorylated proteins at about 120K (Fig. 1d) which could be resolved into a doublet (data not shown). Phorbol ester treatment of neurons increased the phosphorylation of the 120K proteins (Fig. 1d) in a manner similar to that seen with the recombinant receptor expressed in HEK cells. Two-dimensional phosphopeptide maps of tryptic digests of the phosphorylated 120K proteins isolated from cortical neurons (Fig. 2b) were similar to the phosphopeptide maps of the recombinant NR1 subunit. These results show that the 120K proteins are phosphorylated on sites similar to those seen in the recombinant subunit expressed in HEK cells and provide further evidence that the 120K proteins are the NR1 subunit. As in the transfected HEK cells, phosphorylation of all sites was increased by phorbol ester treatment of the neurons (data not shown). The NR1 antibodies also specifically immunoprecipitated a phosphorylated 180K protein which may be the NMDA R2A or 2B subunits (Fig. 1d). These results demonstrate that PKC phosphorylation of the NR1 subunit in neurons is similar, if not identical, to that seen using the recombinant NR1 receptor.

To determine the location of the PKC phosphorylation sites on the NR1 subunit, the phosphorylation of purified bacterial fusion proteins containing different regions of the NR1 subunit was examined and compared to the phosphorylation of the NR1 subunit seen *in situ*. The C-terminal domain of NR1 contains two exons which are alternatively spliced¹⁶ (Fig. 4a), and one of these exons (I) contains several potential sites for PKC phosphorylation based on the characterized consensus sequences for PKC¹⁷. In spite of the fact that the proposed transmembrane topology model of glutamate receptors predict that this region is extracellular^{4,15} (Fig. 4b), we generated a bacterial fusion protein containing the C-terminal 105 amino acids of NR1 (Fig. 4a) and found that it was an excellent substrate for purified PKC *in vitro*. Phosphopeptide mapping using two-dimensional TLC of tryptic digests of the PKC phosphorylated fusion protein (Fig. 2c) demonstrated that the sites of phosphorylation were similar to those phosphorylated by PKC on the NR1 subunit *in situ* in both neurons and HEK cells (Fig. 2). Phosphoamino-acid analysis showed that PKC phosphorylation of the fusion protein occurred mostly on serine residues with a small amount occurring on threonine residues (data not shown).

Taken together, these results suggest that most PKC phosphorylation sites are contained within the C terminal domain of the NR1 subunit and that the phosphorylation of NR1 *in situ* is directly catalysed by PKC. To confirm that the *in situ* phosphorylation occurred on the C-terminal domain and to determine which amino acids were phosphorylated, site-directed mutagenesis was used to eliminate potential phosphorylation sites within the recombinant NR1 subunit. Two different mutants were examined: a deletion mutant which removes the first alternatively spliced exon in the C-terminus (exon I, amino acids 864-900) and produces a subunit identical to the R1C splice variant of the NR1¹⁶ subunit, and another mutant where four serine residues (Ser 889, 890, 896, 897) contained within this exon were mutated to alanine residues (Fig. 4a). PKC phosphorylation of these mutant recombinant receptors expressed in HEK cells was greatly reduced (Fig. 3a), although the level of

* To whom correspondence should be addressed.

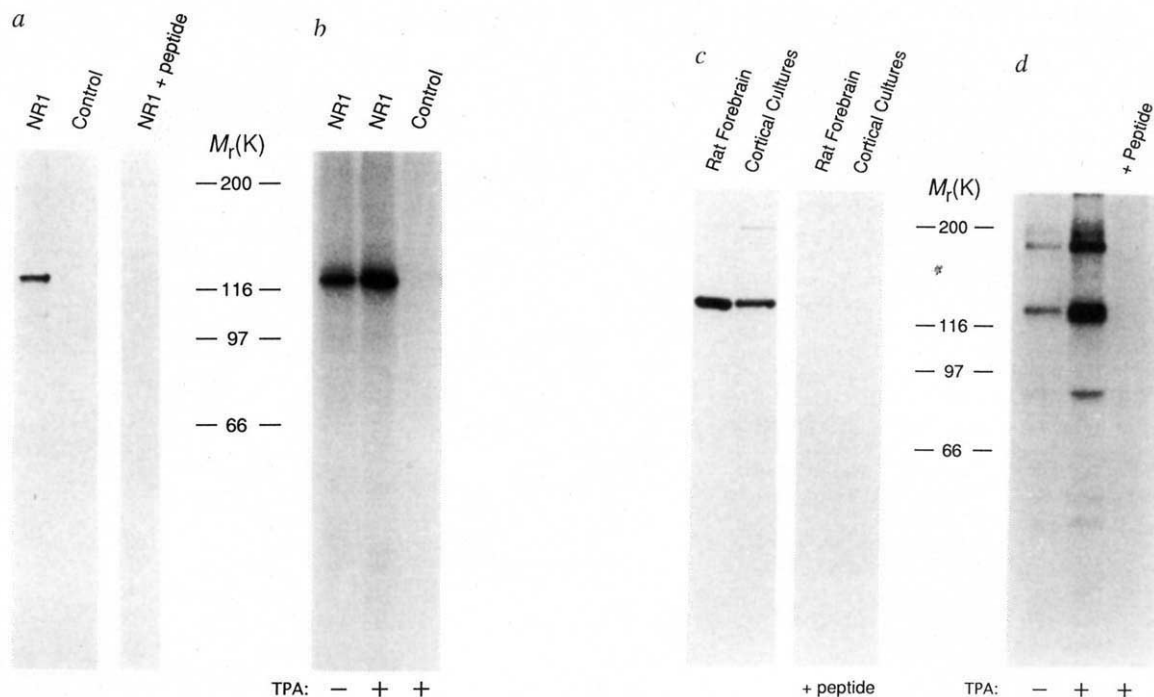
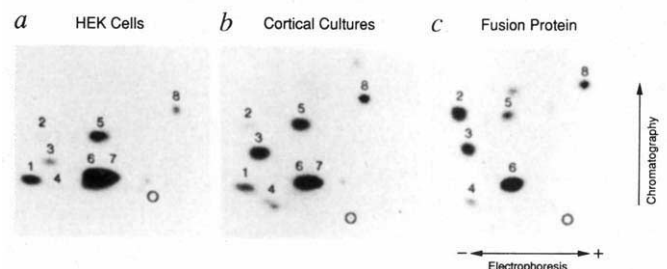


FIG. 1 Phosphorylation of NR1 in transfected HEK cells and in neuronal cell cultures. *a, b*, HEK cells were transiently transfected with a mammalian expression vector containing NR1 cDNA or control DNA, as indicated. *a*, Antibodies recognizing the C-terminal region of NR1 were used to immunoblot cell membranes. Immunodetection by the C-terminal antibody was eliminated by preadsorption of the antibody with the synthetic peptide immunogen (+ peptide) at $15 \mu\text{g ml}^{-1}$. *b*, Transfected HEK cells were prelabelled with [^{32}P]orthophosphate (1 mCi ml^{-1}) for 4 h and then incubated for 30 min in the presence or absence of phorbol ester (100 nM TPA). Cell lysates were immunoprecipitated using a combination of two antibodies which recognize the N and C termini of the NR1 protein. Immunoprecipitates were analysed by SDS-PAGE and autoradiography. *c*, Immunoblots of membranes from adult rat brain and primary cortical cultures using antibodies recognizing the C-terminal region of NMDAR1. Where indicated, antibody was preadsorbed with the synthetic peptide immunogen ($15 \mu\text{g ml}^{-1}$). *d*, Immunoprecipitates of NR1 from cortical cultures prelabelled with [^{32}P] orthophosphate (1 mCi ml^{-1}) for 4 h and then incubated for 30 min in the presence or absence of phorbol ester (100 nM TPA). Where indicated, antibody was preadsorbed with synthetic peptide immunogen ($25 \mu\text{g ml}^{-1}$). **METHODS.** The R1A subunit of rat NMDA cDNA^{5,16} was subcloned into a mammalian expression vector containing the CMV promoter, and

transfected into HEK cells by calcium phosphate coprecipitation²⁵, using $20 \mu\text{g}$ DNA per 10-cm culture dish. Forty hours after transfection, the cells were incubated for four hours in a phosphate-free medium supplemented with 1 mCi ml^{-1} [^{32}P]orthophosphate, and for the last half-hour exposed to either 100 nM TPA, 1% DMSO or 1% DMSO alone. The cells were then solubilized with 1% SDS in cell lysis buffer containing 10 mM sodium phosphate (pH 7.0); 5 mM EDTA; 5 mM EGTA; 100 mM NaCl; 1 mM Na_3VO_4 (ortho); 10 mM sodium pyrophosphate; 50 mM NaF; $10 \mu\text{g ml}^{-1}$ each of leupeptin, antipain, chymostatin and pepstatin; 0.1 mM PMSF; 10 U ml^{-1} Trasylol (aprotinin). Lysates were then diluted 1:5 with cell lysis buffer containing 1% Triton X-100, and sonicated. The NR1 protein was immunoprecipitated using a combination of two affinity-purified antibodies, generated against the synthetic peptides KRAACDPKIVNIGAVLSTRKH and KRAIEREEGQLQCSRHRHS (single-letter code) corresponding to the N and C termini, respectively, of NR1. Neuronal cultures were prepared from 17–18 day fetal rat cerebral cortex and hippocampus as described²⁶. The tissue was dissected, digested with papain and plated at a density of 5×10^6 cells per 6-cm plate. Two to three weeks after plating, cultures were incubated for four hours in a phosphate-free medium supplemented with 1 mCi ml^{-1} [^{32}P]orthophosphate. TPA treatment, cell lysis and immunoprecipitations were done as described above.

FIG. 2 Phosphopeptide map analysis of tryptic phosphopeptides of phosphorylated NR1 protein isolated from transfected HEK cells, neuronal cultures, and a phosphorylated bacterial fusion protein containing the NR1 C terminus. *a*, NR1 was immunoprecipitated from phorbol ester-treated HEK cells, excised from the gel, digested with trypsin and the resulting phosphopeptides analysed by two-dimensional TLC and autoradiography. *b*, Identical phosphopeptide analysis was done on phorbolylated NR1 immunoprecipitated from phorbol ester-treated neurons. *c*, A bacterial fusion protein containing the C-terminal region (amino acids 833–938) of NR1 was phosphorylated *in vitro* with purified PKC and resolved by SDS-PAGE, then excised from the gel and digested with trypsin. The phosphopeptides were resolved by TLC as above. Phosphopeptides with similar migration patterns are indicated by numbers. Open circles represent the origin.

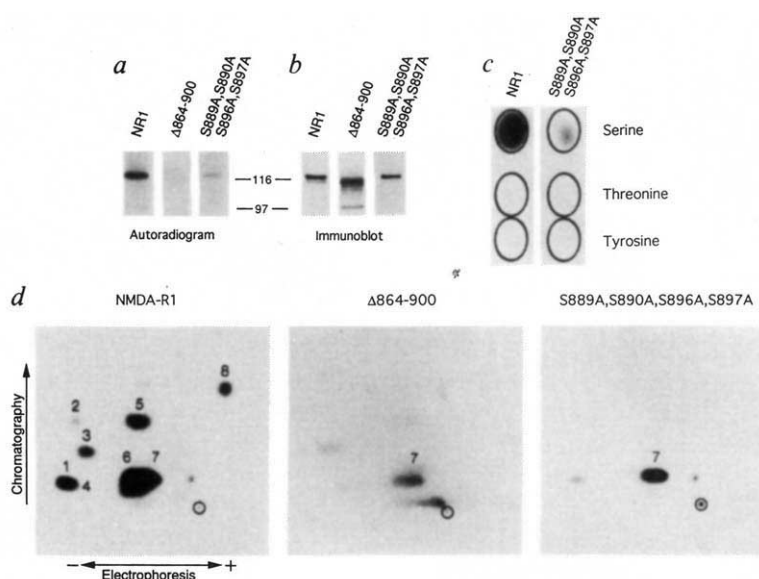
METHODS. The fusion protein of the C terminus of NR1 was generated by amplifying base pairs 2,497–2,814 of the NR1 subunit using polymerase chain reaction (PCR) with the following primers: 5'-ACCATGGAT-CCCGAGATCGCTACAAGCGA-3' and 5'-ACCATGAATTCCTCAGCTCTCCCTATGACG-3', and subcloning the PCR product into pTrcHis vector (Invitrogen). Fusion protein was expressed in BL21 *Escherichia coli*; and purified by chromatography on Ni^{2+} -NTA agarose (Qiagen) in the pres-



ence of 8 M urea. After dialysis against PBS, $1 \mu\text{g}$ fusion protein was incubated with 20 ng purified bovine cardiac PKC for 30 min at 30°C . For phosphopeptide map analysis, phosphoproteins were excised from polyacrylamide gels and cleaved with trypsin (0.3 mg ml^{-1})²⁷. Resulting phosphopeptides were spotted on to TLC plates (Kodak) and subjected to electrophoresis in the horizontal dimension and ascending chromatography in the vertical dimension as described²⁷.

FIG. 3 Phosphorylation of NR1 mutant subunits expressed in HEK cells. Site-directed mutagenesis was used to delete amino-acid residues 864–900 from the C terminus of NR1 and also to convert serines 889, 890, 896 and 897 to alanine residues. These two mutant subunits were compared to wild-type NR1. **a**, Autoradiogram of 32 P-labelled immunoprecipitates of wild-type and mutant NR1 from transfected HEK cells treated with phorbol esters. **b**, Immunoblots of HEK cells transfected with wild-type and mutant NR1 cDNAs. **c**, Phosphoamino-acid analysis of the phosphorylated wild-type and the S889A, S890A, S896A, S897A mutant NR1 subunits shown in **a**. **d**, Phosphopeptide map analysis of 32 P-labelled wild-type and mutant NR1 subunits shown in **a**.

METHODS. The NR1 cDNA was subjected to site-directed mutagenesis as described²⁸. The antisense primer 5'-TCCACCCCGGTGCTCTGCAGGTTCTTCCT-3' was used to generate Δ 864–900. The S889A, S890A, S896A, S897A mutant was created by two rounds of mutagenesis using the antisense primers: 5'-CGTGTCTTTGGCGGCCCTACG-TCTC-3' and 5'-CGTCTCTTGAAGCGCGCGCCAGGG-3'. Mutations were confirmed by DNA sequencing. NR1 mutants were expressed in HEK cells and analysed as described in the legend to Fig 1. For phosphoamino-acid analysis, phosphoproteins were excised from the polyacrylamide gels, digested with trypsin and then hydrolysed in 6 M HCl and analysed by electrophoresis on TLC plates as described²⁹.



expression of the mutant receptors was at least as high as the wild-type receptor as analysed by immunoblots of parallel transfections (Fig. 3b) and by protein staining of the immunoprecipitates (data not shown). Phosphopeptide mapping of the mutant receptors showed that most of the tryptic phosphopeptides were not present; however, one or two major phosphopeptides remained (Fig. 3d). Phosphoamino-acid analysis revealed that the remaining phosphorylation occurred on serine residues with a trace amount on threonine residues (Fig. 3c). Deletion of the 37 amino-acid exon from the C-terminal fusion protein eliminated all of the major sites phosphorylated by PKC (data not shown) suggesting that the residual phosphorylation sites in the mutant NR1 subunits are not on the C-terminal domain.

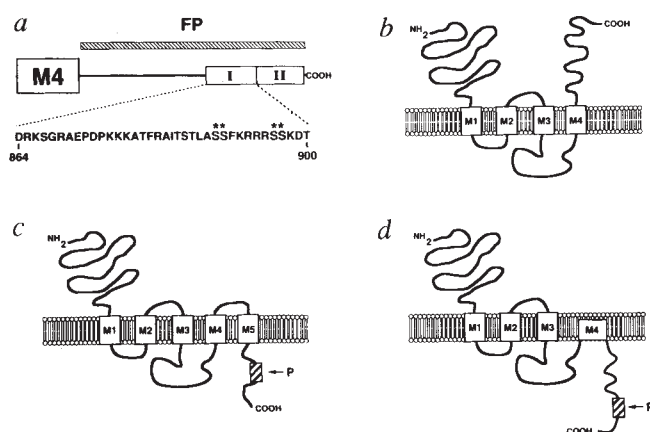


FIG. 4 Structure of the NR1 subunit. **a**, Schematic of the C-terminal region of the NR1 protein showing the position of the two alternatively spliced exons (I and II), the sequence of Exon I containing the serines (S) that were mutated to alanine residues, and the region of the NR1 subunit used to make the bacterial fusion protein (FP). M4, Putative fourth transmembrane domain. **b**, Original proposed transmembrane topology model of the NR1 subunit. **c**, **d**, Two alternative proposed transmembrane topology models to account for the location of the PKC phosphorylation sites. P, Location of the PKC phosphorylation sites.

These results demonstrate that PKC phosphorylates the NR1 subunit of the NMDA receptor on up to four serine residues on a single exon in the C-terminal domain and that alternative splicing of this region regulates this phosphorylation. The C-terminus has been predicted to be on the extracellular side of the membrane based on the transmembrane topology model of the NR1 subunit (Fig. 4b)^{4,5}. This model was based on hydropathy plots of the amino-acid sequence of the NR1 subunit^{4,5} and on the proposed transmembrane topology of subunits for other glutamate receptors^{4,15} and ligand-gated ion channels, such as the nicotinic acetylcholine receptor¹⁸ and the GABA_A receptor¹⁹. But very little independent evidence is available to support this model. Recent data has shown that the non-NMDA glutamate receptor subunit GluR6 is phosphorylated by cAMP-dependent protein kinase on ser 684 which is contained in the predicted major intracellular loop between the putative third and fourth transmembrane domains, confirming the intracellular location of this region²⁰. The data presented here, however, strongly suggests that the C-terminal region of the NR1 subunit is intracellular, because it is accessible to PKC, and it seems unlikely that PKC phosphorylation occurs on an extracellular domain of the NR1 protein. In fact, recent immunocytochemical studies using C-terminal antibodies are consistent with an intracellular localization of the C terminus^{21,23}. One possible model to reconcile these results is that the C-terminal domain crosses the membrane with a fifth transmembrane region (Fig. 4c), in spite of the fact that there are no obvious hydrophobic regions between the proposed fourth transmembrane domain and the phosphorylation sites. Alternatively, the fourth transmembrane domain may not cross the membrane, leaving the C-terminus on the intracellular side of the membrane (Fig. 4d). Whether these results are relevant to the transmembrane topology of other ligand-gated ion channels is not clear.

Previous studies have reported that activation of PKC can either potentiate^{7,11} or inhibit¹², NMDA receptor function. Our results demonstrate that PKC directly phosphorylates the NR1 subunit of NMDA receptors. Moreover, we show that alternative splicing controls PKC phosphorylation of the NR1 subunit and may therefore regulate the functional modulation of NMDA receptors by PKC. Protein phosphorylation of ligand-gated ion channels appears to be a central mechanism for regulating synaptic efficacy^{13,24}. Alternative splicing of receptor subunits, com-

bined with the wide variety of receptor subunits and protein kinases, increases the potential diversity of short- and long-term modulation of synaptic transmission through protein phosphorylation. □

Received 25 March; accepted 28 May 1993.

1. Bliss, T. V. P. & Collingridge, G. L. *Nature* **361**, 31–39 (1993).
2. Choi, D. W. & Rothman, S. M. *A. Rev. Neurosci.* **13**, 171–182 (1990).
3. Constantine-Paton, M., Cline, H. T. & Debski, E. A. *Rev. Neurosci.* **13**, 129–154 (1990).
4. Nakanishi, S. *Science* **258**, 597–603 (1992).
5. Moriyoshi, K. et al. *Nature* **354**, 31–37 (1991).
6. Monyer, H. et al. *Science* **256**, 1217–1221 (1992).
7. Kutsuwada, T. et al. *Nature* **358**, 36–41 (1992).
8. Chen, L. & Huang, L.-Y. *Nature* **356**, 521–523 (1992).
9. Kelso, S. R., Nelson, T. E. & Leonard, J. P. *J. Physiol.* **449**, 705–718 (1991).
10. Durand, G. M. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 1–5 (1992).
11. Urushihara, H., Tohda, M. & Nomura, Y. *J. Biol. Chem.* **267**, 11697–11700 (1992).
12. Markham, H. & Segal, M. *J. Physiol.* **457**, 491–501 (1992).
13. Raymond, L. R., Blackstone, C. D. & Huganir, R. L. *Trends Neurosci.* **16**, 147–153 (1993).
14. Chen, L. & Huang, L.-Y. *Neuron* **7**, 319–326 (1991).
15. Gasic, G. P. & Hollmann, M. A. *Rev. Physiol.* **54**, 507–536 (1992).
16. Sugihara, H., Moriyoshi, K., Ishii, T., Masu, M. & Nakanishi, S. *Biochem. Biophys. Res. Commun.* **185**, 826–832 (1992).

17. Kennelly, P. J. & Krebs, E. G. *J. Biol. Chem.* **266**, 15555–15558 (1991).
18. Claudio, T., Ballivet, M., Patrick, J. & Heinemann, S. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1111–1115 (1983).
19. Schofield, P. R. et al. *Nature* **326**, 221–227 (1987).
20. Raymond, L. R., Blackstone, C. D. & Huganir, R. L. *Nature* **361**, 637–641 (1993).
21. Petralia, R. S. & Wenthold, R. J. *J. comp. Neurol.* **318**, 329–354 (1992).
22. Martin, L. J., Blackstone, C. D., Levey, A. I., Huganir, R. L. & Price, D. L. *Neuroscience* **53**, 327–358 (1993).
23. Molnár, E. et al. *Neuroscience* **53**, 307–326 (1993).
24. Swope, S. L., Moss, S. J., Blackstone, C. D. & Huganir, R. L. *FASEB J.* **6**, 2514–2523 (1992).
25. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* 16.37–16.38 (Cold Spring Harbor Laboratory Press, New York, 1989).
26. Murphy, T. H. & Baraban, J. M. *Dev. Brain Res.* **57**, 146–150 (1990).
27. Huganir, R. L., Miles, K. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6968–6972 (1984).
28. Kunkel, T. A., Roberts, J. D. & Zabour, D. L. *Meth. Enzym.* **154**, 367–382 (1987).
29. Hirano, A. A., Greengard, P. & Huganir, R. L. *J. Neurochem.* **50**, 1447–1455 (1988).

ACKNOWLEDGEMENTS. We thank C. Blackstone and L. Raymond for critical reading of this manuscript, C. Finch for help in the preparation of the manuscript, R. S. Fiore and T. Murphy for help with the primary cortical cultures and C. Doherty and L. Moritz for technical assistance. We thank S. Nakanishi (Kyoto University) for the cDNA clones of NMDAR1. Supported by the Howard Hughes Medical Institute (R.L.H.) and an MSTP Fellowship (W.T.).

Severe atherosclerosis in transgenic mice expressing simian cholesteryl ester transfer protein

Keith R. Marotti, Christine K. Castle, Timothy P. Boyle, Alice H. Lin, Robert W. Murray & George W. Melchior

Molecular Biology Research and Metabolic Diseases Research, Upjohn Laboratories, 7242-209-7 Kalamazoo, Michigan 49001, USA

CHOLESTERYL ester transfer protein (CETP) is a plasma protein that mediates the exchange of neutral lipids among the lipoproteins^{1–3}. Because the principal core lipid of very-low-density lipoprotein (VLDL) is triglyceride and that of high-density lipoprotein (HDL) is cholesterol ester, CETP mediates a 'heteroexchange' of cholesterol ester for triglyceride between those lipoproteins. As a result, animals that express CETP tend to have higher VLDL and low-density lipoprotein (LDL) cholesterol levels, whereas those with no CETP activity tend to have high HDL cholesterol levels². Because VLDL and LDL are associated with the progression of atherosclerosis, and HDL are considered anti-atherogenic, CETP could be an 'atherogenic' protein, that is, given the other conditions required for atherosclerosis to develop, expression of CETP would accelerate the rate at which the arterial lesions progress. We report here that transgenic mice expressing CETP had much worse atherosclerosis than did non-expressing controls, and we suggest that the increase in lesion severity was due largely to CETP-induced alterations in the lipoprotein profile.

We chose as our model of atherosclerosis the cholesterol-fed, C57BL/6 mouse, an inbred strain with well defined genetics and in which the development of atherosclerosis has been well characterized^{4–6}. Atherosclerotic lesions consistently developed in C57BL/6 mice at a defined location in the proximal aorta within 12–18 weeks after the mice began consuming an atherogenic diet, and a method for quantifying the lesions was described. To evaluate the effect of CETP on the development of atherosclerosis, we produced transgenic C57BL/6 mice that express cynomolgus monkey CETP at various levels⁷. We used only males for these initial studies, because males are significantly more resistant than females to development of the diet-induced disease^{5,8} and, if CETP expression did have an

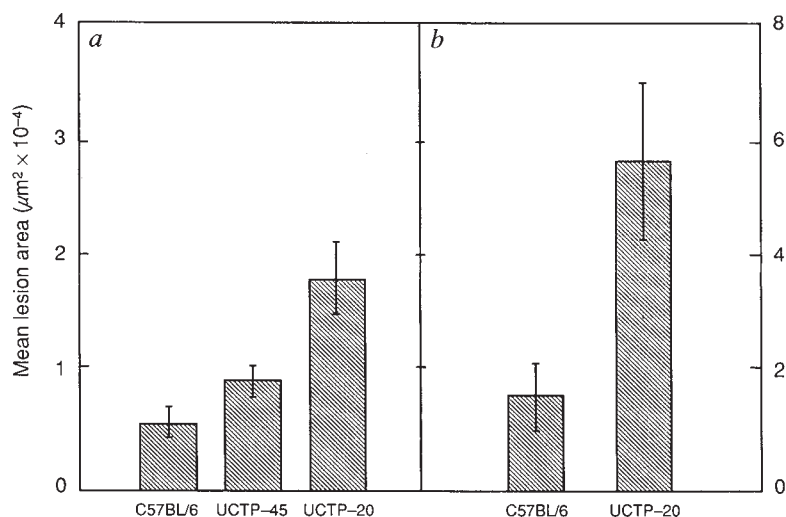
effect on development of the disease, the effect might be more evident in a model that tended to be resistant. Three strains of mice were used for these studies: C57BL/6 mice, which have little or no CETP activity; UCTP-45, a transgenic C57BL/6 mouse expressing simian CETP at relatively low levels; and UCTP-20, a transgenic C57BL/6 mouse expressing CETP at relatively high levels (Table 1). The mice ranged in age between 6 and 9 months and were fed a high-fat, high-cholesterol semi-synthetic diet for either 18 weeks or 28 weeks, at which point they were killed, and their hearts removed, fixed and sectioned. The sections were processed and the lesions quantified essentially as in ref. 5.

Table 1 shows the effect of CETP expression on the plasma cholesterol concentration and distribution and the plasma apolipoprotein A-I and apolipoprotein B levels of the mice before and after they began consuming an atherogenic diet. Note that the VLDL+LDL cholesterol and apo B levels were slightly higher in the CETP transgenic mice than in the C57BL/6 controls when they were consuming the chow diet ($t=0$; Table 1) and that the VLDL+LDL cholesterol and apo B levels increased to a greater extent in the CETP mice than in the C57BL/6 mice when they were fed the atherogenic diet. The CETP transgenic mice also had significantly lower HDL-cholesterol and apo A-I levels than the C57BL/6 mice when they were consuming the chow diet and, although the HDL-cholesterol and apo A-I levels in the CETP mice increased in response to the atherogenic diet, they failed to reach even the basal HDL levels of the C57BL/6 controls. Statistical analysis confirmed that there was a significant diet – gene interaction on (VLDL+LDL) and HDL levels ($P<0.01$ in both cases). Thus, expression of CETP in the C57BL/6 males caused what would be considered an atherogenic shift in the (VLDL+LDL)/HDL ratio, and the change in that ratio was most severe in the CETP mice fed an atherogenic diet.

Figure 1a shows the mean lesion area per mouse in C57BL/6 mice and CETP-transgenic mice that had consumed the atherogenic diet for 18 weeks, and Fig. 1b shows the lesion area in mice that had consumed the diet for 28 weeks. The effect of CETP expression on the severity of the disease at each time point is apparent, and that effect was statistically significant in both instances. Figure 2 shows lesions typical of those seen in UCTP-20 and C57BL/6 mice at 18 and 28 weeks. Note that the lesions in the C57BL/6 mice, besides being less severe, tended to be more diffuse, appearing as a thin ribbon of sudanophilia along the intimal surface. By contrast, the lesions in the CETP mice were large, focal lesions, more typical of those that form in rabbits and monkeys fed atherogenic diets. None of the

FIG. 1 Mean ($\pm$ s.e.m.) atherosclerotic lesion area in C57BL/6 mice and CETP-transgenic mice that had consumed the atherogenic diet (see Table 1) for 18 weeks (Fig. 1a) or 28 weeks (Fig. 1b). Note the difference in the ordinate scale. Statistical analysis of the 18-week data (one-way analysis of variance, log-transformation) indicated that the probability of obtaining a greater t was: C57BL/6 versus UCTP-45, 0.181; C57BL/6 versus UCTP-20, 0.001; UCTP-45 versus UCTP-20, 0.043. Statistical analysis (independent sample t -test, reciprocal transformation) of the 28 week data (C57BL/6 versus UCTP-20) indicated the probability of obtaining a greater t was 0.006.

METHODS. At the end of each study, the mice were anaesthetised and exsanguinated through the brachial artery. Their hearts were removed, fixed (4% phosphate-buffered formaldehyde), embedded (25% gelatin) and sectioned exactly as described in ref. 5. Briefly, 10- μ m cross-sections were taken sequentially, starting just above the aortic valve and moving along the aorta in the direction of blood flow, until 40 sections had been taken. The sections were processed and the lesion area of every eighth section was quantified as described⁵. The sum of those areas for a given mouse was taken as the numerical value that best characterized the lesion severity in that mouse. The mean $\pm$ s.e.m. of those values for each study group (Table 1) are shown.



C57BL/6 mice exhibited the large focal lesions seen in the CETP mice. Thus, it appears that CETP expression not only increased the severity of the atherosclerosis produced at 18 and 28 weeks, but also had a profound effect on the morphology of the lesion.

Although the progression of atherosclerosis is a complex process, influenced by a multiplicity of factors in man, its aetiology is somewhat more well defined in animal models of the disease. In almost every case, a marked increase in the concentration of VLDL+LDL cholesterol levels is required to produce the disease, and in most instances that change is accompanied by a decrease in the plasma HDL-cholesterol levels⁹. To determine the extent to which the CETP-induced increase in lesion severity was due to the change in the lipoprotein profile, lesion severity at 18 weeks was regressed on the (VLDL+LDL)/HDL cholesterol ratio. The results of that

analysis are shown in Fig. 3. As is apparent, a strong positive correlation existed. An almost identical relationship existed at 28 weeks (not shown). Note that the y-intercept of the regression line was 0.42. That could be interpreted as an indication that mice with a (VLDL+LDL)/HDL ratio greater than 0.42 will develop some lesions after 18 weeks; however, the mean (VLDL+LDL)/HDL ratio of a separate group of 14 chow-fed, UCTP-20 mice averaged 0.65 ± 0.07 for well over 28 weeks, and 5 of those mice had (VLDL+LDL)/HDL ratios that averaged 0.8 or higher during that period. Yet, when the aortae of those mice were examined, no arterial lesions were evident. We interpret this as indication that a degree of hypercholesterolaemia is required in addition to the altered (VLDL+LDL)/HDL ratio to produce lesions within 28 weeks. It remains to be determined whether, over longer periods, any of the CETP mice

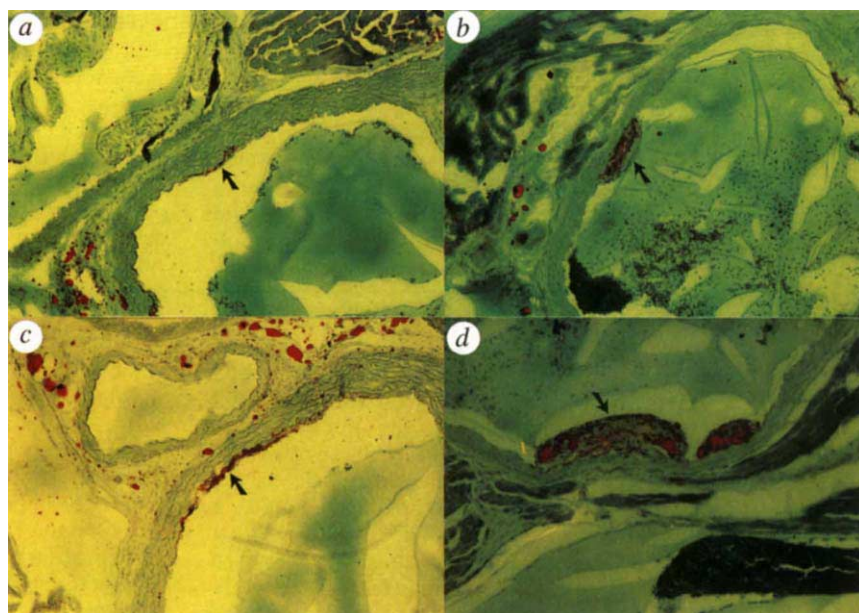


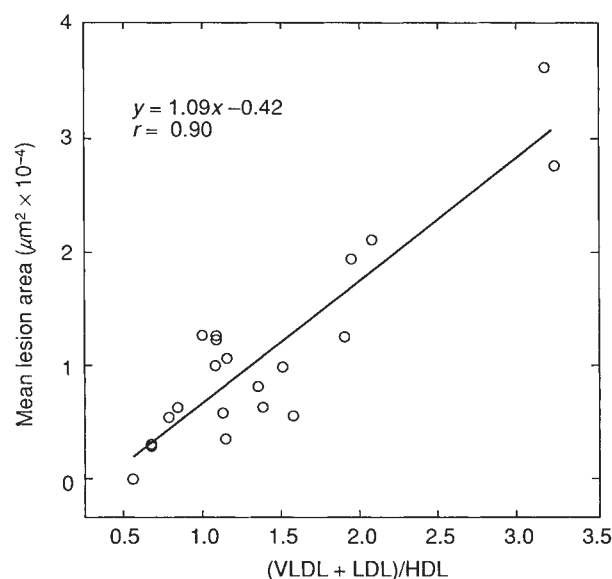
FIG. 2 Cross-sections of the proximal aorta from control and CETP-transgenic mice showing the effect of CETP expression on lesion size and morphology. The sections were processed as described in Fig. 1, stained with Oil Red O and haematoxylin and counter-stained with light green and are typical of those seen in each study group. a, C57BL/6 mice that had consumed the atherogenic diet for 18 weeks; b, UCTP-20 mice that had consumed the diet for 18 weeks; c, C57BL/6 mice that had consumed the diet for 28 weeks; d, UCTP-20 mice that had consumed the diet for 28 weeks. Magnification, $\times 46$.

TABLE 1 Plasma cholesterol and apolipoprotein A-I and B concentrations in C57BL/6 and CETP transgenic mice fed an atherogenic diet

Time (weeks)	Mouse strain	N	Total cholesterol	VLDL+LDL cholesterol	HDL cholesterol	(VLDL+LDL)/HDL ratio	Apo A-I	Apo B
0	C57BL/6	15	96±4	14±1	82±4	0.18±0.02	130±4	14±1
0	UCTP-45	5	67±5	16±1	46±3	0.35±0.04	89±9	20±1
0	UCTP-20	16	61±3	22±3	39±3	0.62±0.08	90±4	25±1
18	C57BL/6	8	176±3	75±8	88±3	0.85±0.08	190±7	26±1
18	UCTP-45	5	167±7	94±5	73±5	1.31±0.13	116±11	30±2
18	UCTP-20	9	173±14	98±9	55±4	1.90±0.29	125±5	34±2
28	C57BL/6	7	189±6	81±3	96±3	0.85±0.04	196±7	25±1
28	UCTP-20	7	189±10	130±16	61±3	2.18±0.32	125±7	33±2

The study was initiated by changing the diets of the mice from Purina rodent chow ($t=0$) to a semi-synthetic diet containing 1.0% cholesterol, 15% fat from dairy butter, 50% sucrose, 20% casein and 0.5% sodium cholate in addition to cellulose, vitamins and minerals¹⁰. The mice were maintained on those diets for 18 or 28 weeks, and blood samples taken from the orbital sinuses of all mice in the study at $t=0$ and at 7 week intervals thereafter. The mice in the 18 week group also had blood samples taken at 18 weeks. The value at $t=0$ are those from samples taken just before the mice were switched to the semi-synthetic diet. The lipid and apolipoprotein values were determined as described previously¹¹ and are expressed as mg dl⁻¹ of plasma. To obtain a single number which best described a lipid or apolipoprotein concentration of a given mouse over the entire study, the values obtained for that mouse at each bleeding were averaged. The value in the table represents the mean $\pm$ s.e.m. of those numbers for all of the mice in a given group. The plasma CETP activity (arbitrary units, mean $\pm$ s.e.m.) for each of the 3 strains of mice was: C57BL/6 0.004 ± 0.001 AU; UCTP-45, 2.6 ± 0.2 AU; and UCTP-20, 6.7 ± 0.4 AU; where one AU is the activity in an equal volume of cynomolgus monkey plasma standard¹². Statistical summary: One-way analysis of variance (ANOVA) using our entire population of chow-fed UCTP-20 and UCTP-45 mice and 15 chow-fed C57BL/6 mice showed that there was a statistically significant ($P < 0.05$) effect of CETP expression on all of the lipoprotein profile parameters shown (both UCTP-20 and UCTP-45 mice were different from C57BL/6 mice). When analyses were limited to these data, the VLDL+LDL cholesterol and the (VLDL+LDL)/HDL ratio of UCTP-45 mice were not statistically different from C57BL/6 mice; all of the other measurements were. Three-way ANOVA (strain, diet and time) showed that the diet had a statistically significant effect on all of the parameters measured, regardless of strain, and that, with the exception of total plasma cholesterol, the UCTP-20 and UCTP-45 mice responded differently to the diet than did the C57BL/6 mice (there was not a significant effect of CETP expression on total plasma cholesterol of mice consuming the semi-synthetic diet). Finally, most of the parameters measured appeared to have reached a 'steady-state' by 14 weeks, and there was not a statistically significant net change over 28 weeks.

FIG. 3 Statistical relationship between mean lesion area and (VLDL+LDL)/HDL ratio at 18 weeks. The mean lesion area for each mouse was plotted against the mean (VLDL+LDL)/HDL cholesterol ratio (determined as described under Table 1). The equation produced by a least-squares fit of the data is shown in the upper left-hand corner, as is the correlation coefficient. These data indicated that 81% of the change in lesion area could be accounted for by change in the mean (VLDL+LDL)/HDL cholesterol ratio.



will develop lesions spontaneously.

The high degree of correlation between lesion area and the (VLDL+LDL)/HDL ratio in the hypercholesterolaemic mice, and the fact that the total plasma cholesterol concentrations were not significantly different among those groups (Table 1), indicates that lesion progression in that model was not so much a function of plasma cholesterol concentration, as it was a

function of how that cholesterol was partitioned between the lipoproteins. We interpret these data as an indication that the CETP-induced alteration in cholesterol distribution was the principal reason that the arterial lesions developed more rapidly in the CETP transgenic mice, and therefore conclude that CETP is, in fact, an 'atherogenic' protein. □

Received 3 February; accepted 8 April 1993.

1. Zilversmit, D. B., Hughes, L. B. & Balmer, J. *Biochim. biophys. Acta* **409**, 393–398 (1975).
2. Rye, K. A. & Barter, P. J. *Structure and Function of Apolipoproteins* (ed. Rosseneu, M.) 401–426 (CRC Press, Boca Raton, 1992).
3. Tall, A. R. *J. Int. Med.* **231**, 661–668 (1992).
4. Paigen, B., Morrow, A., Brandon, C., Mitchell, D. & Holmes, P. *Atherosclerosis* **57**, 65–73 (1985).
5. Paigen, B., Morrow, A., Holmes, P. A., Mitchell, D. & Williams, R. A. *Atherosclerosis* **68**, 231–240 (1987).

6. Paigen, B. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 3763–3767 (1987).
7. Marotti, K. R. *et al. Arterioscler. Thromb.* **12**, 736–744 (1992).
8. Paigen, B., Holmes, P. A., Mitchell, D. & Albee, D. *Atherosclerosis* **64**, 215–221 (1987).
9. Mahley, R. W. *J. Med. Clin. N. Am.* **66**, 375–402 (1982).
10. Nishina, P. M., Verstuyt, J. & Paigen, B. *J. Lipid Res.* **31**, 859–869 (1990).
11. Castle, C. K., Colca, J. R. & Melchior, G. W. *Arterioscler. Thromb.* **13**, 302–309 (1993).
12. Pape, M. E., Rehberg, E. F., Marotti, K. R. & Melchior, G. W. *Arterioscler. Thromb.* **11**, 1759–1771 (1991).

ACKNOWLEDGEMENTS. We thank B. Paigen, Jackson Laboratory, Bar Harbor, Maine for helpful discussions.

Cyclic ADP-ribose as an endogenous regulator of the non-skeletal type ryanodine receptor Ca^{2+} channel

László G. Mészáros*, Judit Bak* & Alice Chu†

*Department of Physiology and Endocrinology, Medical College of Georgia, Augusta, Georgia 30912, USA

†Cardiovascular Sciences, Department of Medicine, Baylor College of Medicine, Houston, Texas 77030, USA

THE skeletal and cardiac isoforms¹ of the ryanodine receptor Ca^{2+} channel (RyRC) constitute the Ca^{2+} release pathway in sarcoplasmic reticulum of skeletal and cardiac muscles, respectively^{2,3}. A direct mechanical and a Ca^{2+} -triggered mechanism (Ca^{2+} -induced Ca^{2+} release) have been respectively proposed to explain the *in situ* activation of Ca^{2+} release in skeletal and cardiac muscle^{4,5}. In non-muscle cells, however, where the RyRC also participates in Ca^{2+} signalling⁶⁻¹¹, the mechanism of RyRC activation is unknown. Cyclic adenosine 5'-diphosphoribose (cADPR)¹², which is present in many mammalian tissues¹³, has been reported to induce Ca^{2+} release from ryanodine-sensitive intracellular Ca^{2+} stores in sea urchin eggs¹⁴. Here we provide evidence that cADPR directly activates the cardiac but not the skeletal isoform of the RyRC. This, together with results on sea urchin eggs¹⁴, suggests that cADPR is an endogenous activator of the non-skeletal type of RyRC and may thus have a role similar to inositol 1,4,5-trisphosphate¹⁵ in Ca^{2+} signalling.

Using metallochromic calcium indicators, it has been shown¹⁴ that, after ATP-dependent Ca^{2+} accumulation in sea urchin egg homogenates, the addition of cADPR induces a sizeable Ca^{2+} release, which is apparently sensitive to modulators of the RyRC¹⁶. In similar experiments, we have found that cADPR can induce Ca^{2+} release from cardiac, but not from skeletal sarcoplasmic reticulum (SR) (not shown), whereas caffeine readily elicits Ca^{2+} release from both SR vesicles¹⁷. We also assessed the effect of cADPR on Ca^{2+} release from both cardiac and skeletal SR vesicles loaded passively with $^{45}\text{Ca}^{2+}$, a protocol that may provide a better control over extra- and intravesicular Ca^{2+} concentrations. At 10^{-7} M free extravesicular Ca^{2+} concentration ($[\text{Ca}^{2+}]$), cADPR significantly enhanced the rate of $^{45}\text{Ca}^{2+}$ efflux from cardiac, but not from skeletal vesicles (Fig. 1a). Furthermore, cADPR also increased the rate of $^{45}\text{Ca}^{2+}$ efflux from brain microsomes (Fig. 1b), which show ryanodine-sensitive Ca^{2+} release⁹ and express the cardiac (instead of the skeletal) isoform of the RyRC (ref. 1). The cADPR concentrations for half-maximal stimulation of Ca^{2+} release from cardiac SR and brain microsomes were similar (0.15 and 0.11 μM , respectively).

Ca^{2+} release activated by cADPR from both cardiac SR and brain microsomes (Fig. 1) was inhibited by 50 μM ryanodine, which, at this high concentration, blocks the RyRC¹⁸. This suggests that the effect of cADPR on Ca^{2+} efflux stems from its interaction with the RyRC. To test this hypothesis further, we also examined the influence of cADPR on the single channel properties of the cardiac RyRC incorporated into planar lipid bilayers (Fig. 2). Unitary channel conductance was first recorded in the presence of 10^{-5} M cytoplasmic (*cis*) $[\text{Ca}^{2+}]$ (traces in Fig. 2a). Long-lived as well as frequent but short-lived open states were observed, with rare silent periods. When EGTA was added to the *cis* chamber to lower the $[\text{Ca}^{2+}]$ to 10^{-8} M, current fluctuations became rarer (traces in Fig. 2b) and, when present, they were grouped into clusters separated by long silent periods. But the subsequent addition of micromolar cADPR to the *cis*

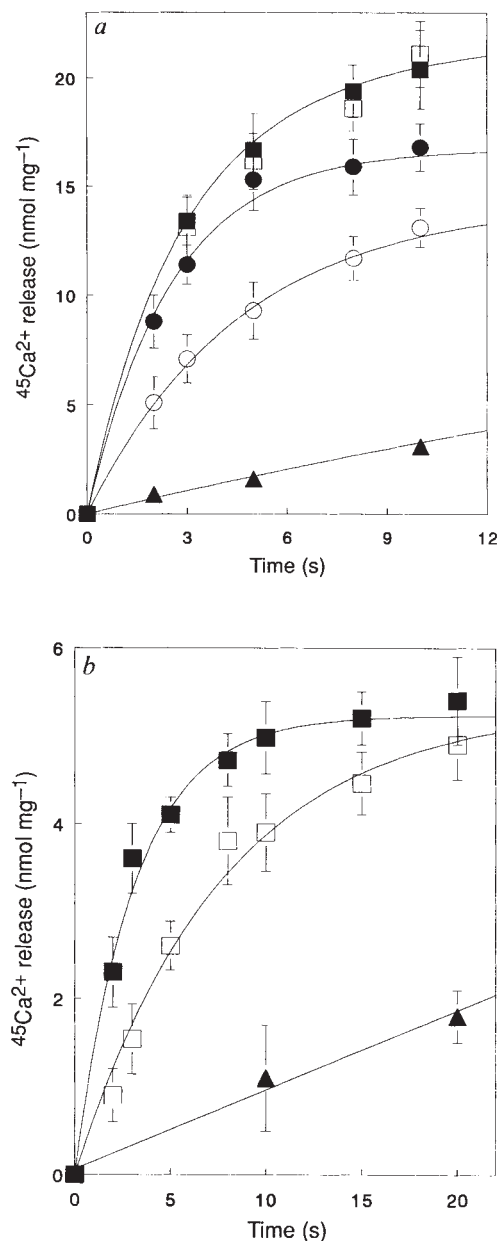
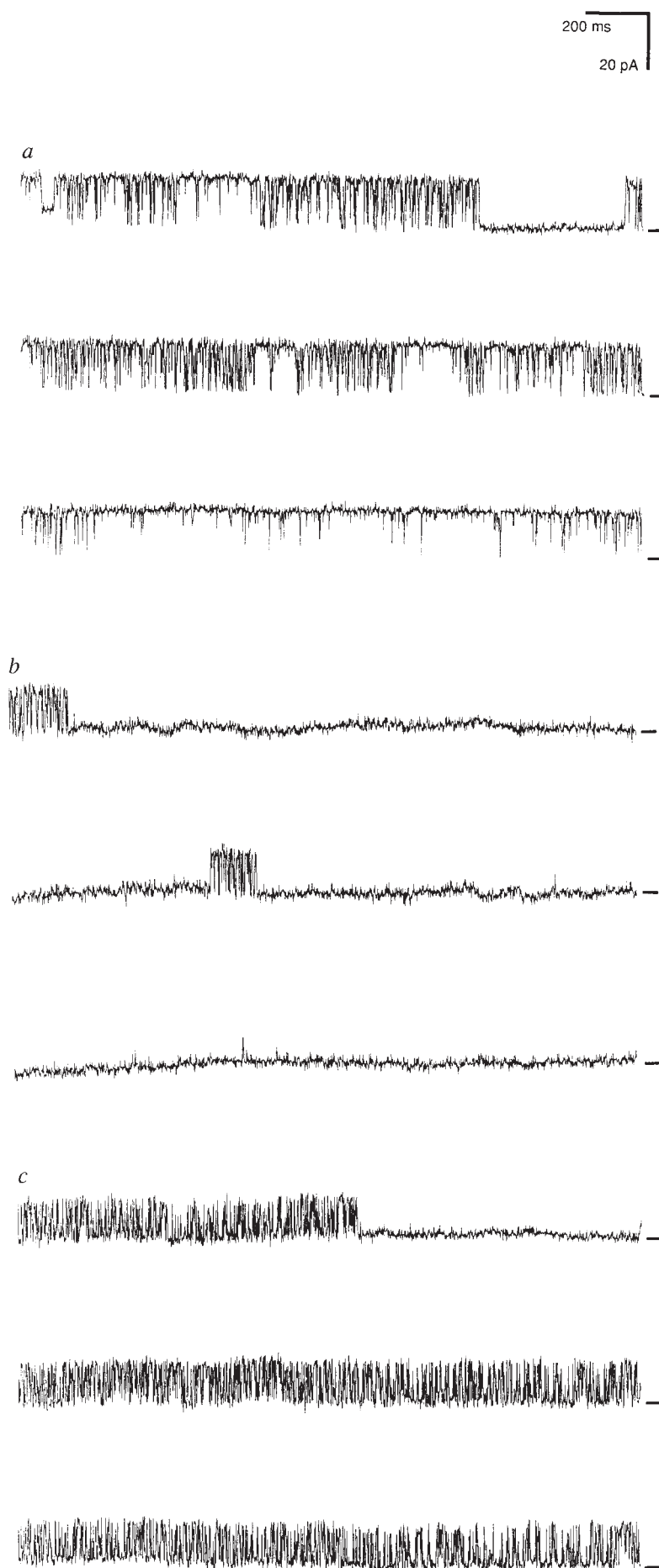


FIG. 1 The effect of cADPR on $^{45}\text{Ca}^{2+}$ efflux from cardiac and skeletal SR and brain microsomes. Skeletal¹⁷ and cardiac²⁴ SR and brain microsomes²⁵ were prepared as described. The final pellets were resuspended in 0.3 M sucrose, 20 mM 3-(N-morpholino)-propanesulphonic acid (MOPS) (pH 6.8) and a mixture of protease inhibitors (0.1 mM phenylmethylsulphonyl fluoride, 10 $\mu\text{g ml}^{-1}$ aprotinin, 1 mM benzamide, 1 $\mu\text{g ml}^{-1}$ leupeptin and 2 $\mu\text{g ml}^{-1}$ pepstatin A). The vesicles (4 to 5 mg ml^{-1}) from cardiac (circles and triangles), skeletal (squares) SR (a) and brain microsomes (b) were preincubated overnight on ice in 2 mM $^{45}\text{CaCl}_2$, 150 mM KCl and 20 mM MOPS (pH 6.8). They were then diluted 20-fold into a Ca^{2+} -release medium containing 150 mM KCl, 20 mM MOPS (pH 6.8) and 250 μM EGTA to adjust the pCa to 7. $^{45}\text{Ca}^{2+}$ efflux was quenched with ice-cold quench solution containing 1 mM LaCl_3 , 10 mM MgCl_2 , 150 mM KCl, and 20 mM MOPS (pH 6.8)⁹. After Millipore filtration and washing the filters with the quench solution the radioactivity retained by the filter was determined by liquid scintillation counting. Open symbols represent controls (no additions), and filled symbols represent the addition of either 1 μM cADPR (squares and circles) or 1 μM cADPR and 50 μM ryanodine (triangles; also present during calcium loading). Data represent the average of measurements carried out with 5 cardiac and 2 skeletal SR and 3 brain microsomal preparations. Error bars are $\pm$ s.d. Solid lines are the best fit to the data.

FIG. 2 The activation of single RyRCs by cADPR. Channel conductance was recorded at -30 mV, in a 400 mM symmetric (*cis/trans*) caesium methanesulphonate–20 mM caesium HEPES (pH 7.2)^{26–28} in sequential order from top to bottom (see Methods). The channel openings are shown as upward deflections, with the flat baseline as the closed position (tick marks on the right). *a*, Traces were recorded in the presence of contaminating Ca^{2+} (8–13 μM , as determined by atomic absorption spectrometry; $\text{pCa}=5$). *b*, then 150 μM EGTA was added to the *cis* chamber to lower the $[\text{Ca}^{2+}]$ to $\text{pCa}=8$. *c*, Traces were subsequently recorded after adding 1 μM *cis* cADPR at $\text{pCa}=8$. The Cs^+ conductance of the channel was 456 pS.

METHODS. A planar lipid bilayer was formed across an aperture (100 μm) in a Delrin cup from a mixture of (50 mg ml^{-1}) palmitoyl oleoyl phosphatidylethanolamine and palmitoyl oleoyl phosphatidylcholine (Avanti Polar Lipids) (7:3, v/v) in decane^{26–28}. Cardiac SR vesicles (5–20 $\mu\text{g ml}^{-1}$) were added to the *cis* chamber, which corresponds to the cytoplasmic side of the SR membrane. The *trans* chamber is referred to ground, and is the SR luminal side. Channels were identified by their typical response to micromolar ryanodine, which is to lock the channel in a subconducting state²⁹. A pulse protocol of -50 to $+50$ mV was applied to incorporate channels from the vesicles into the bilayer, in a medium containing 400 mM/200 mM (*cis/trans*) caesium methanesulphonate–20 mM HEPES (pH 7.2) and contaminating Ca^{2+} . After fusion, the gradient across the bilayer was dissipated (see above), and single channel activity was recorded in 32-s segments. Sampling frequency was 2.5 kHz; data were filtered at 1 kHz with an 8-pole bessel filter. For data acquisition, software and hardware (pClamp, TI-1 interface, Axon Instruments) were interfaced with a 386 microcomputer. pClamp software was used to analyse data.



chamber distinctly raised the frequency of channel opening (traces in Fig. 2c). The records in the presence of cADPR appeared more uniform: fast gating, with short-lived open states and occasional silent periods. The mean-time constants of biexponential fits to open and closed lifetime histograms binned from traces *a* and *c* (Fig. 2) also show different characteristics of gating in the presence of micromolar extravesicular Ca^{2+} or cADPR (Table 1). In the presence of cADPR, open times were shorter and closed times were longer than in micromolar Ca^{2+} . Table 1 also summarizes the pooled open probability (P_o) values computed from 400–600 s of recording time for each condition:

TABLE 1 Effect of cADPR on single channel properties of cardiac RyRC

	Open time (ms)		Closed time (ms)		P_o
	t_1	t_2	t_1	t_2	
pCa=5, control	1.53	9.70	0.79	3.42	0.64 ± 0.15
pCa=8, control	—	—	—	—	0.09 ± 0.05
pCa=8, cADPR	0.56	1.04	2.14	8.62	0.35 ± 0.08

The opening probabilities (P_o) were computed by setting the threshold at half amplitudes²³. The differences between groups (pCa=5, control, $n=5$ channels; pCa=8, control, $n=5$ channels; pCa=8, with 1 μM cADPR, $n=5$ channels) were analysed by variance using the program Instat (Graphpad, San Diego) and found to be significant ($P < 0.05$). Data are given as the mean $\pm$ s.d. The mean lifetime constants (t_1 and t_2) were computed from the traces in Fig. 2 *a* and *c* by fitting biexponentials to open and closed time histograms obtained by linear binning.

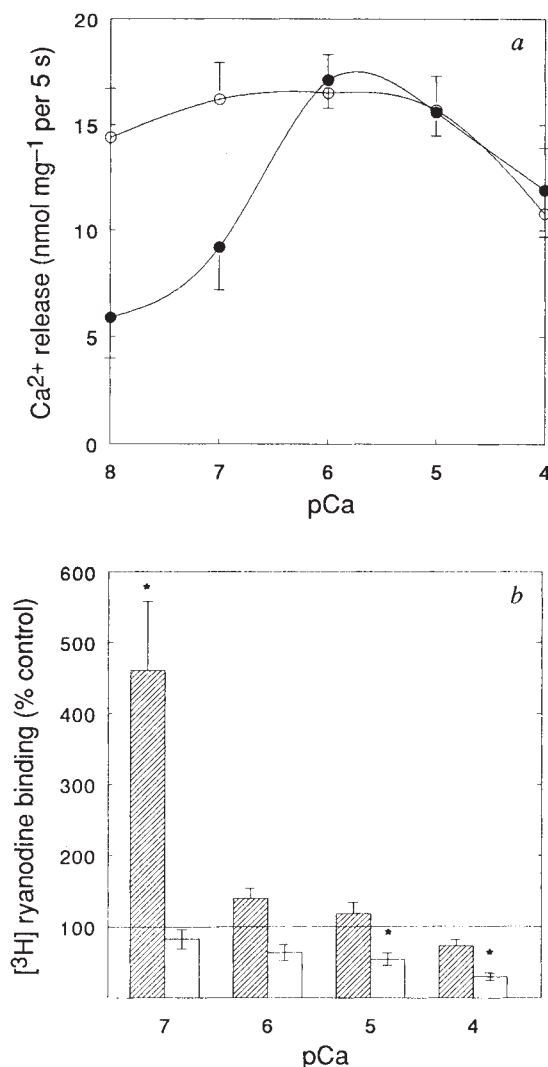


FIG. 3 Ca^{2+} -dependent effect of cADPR on *a*, $^{45}\text{Ca}^{2+}$ efflux, and *b*, $[^3\text{H}]$ ryanodine binding in cardiac and skeletal SR. *a*, The rate of $^{45}\text{Ca}^{2+}$ efflux was determined at 5 s after dilution into releasing media in the absence (filled circles) and in the presence (open circles) of cADPR (1 μM) as described in Fig. 1 legend, except that the concentration of EGTA was varied to obtain the pCa indicated. Error bars are $\pm$ s.d. from 4 to 6 determinations. *b*, $[^3\text{H}]$ ryanodine binding was determined as described¹⁹. Cardiac (hatched bars) and skeletal (open bars) SR membranes (0.15 mg ml^{-1}) were incubated in a medium containing 5 nM $[^3\text{H}]$ ryanodine, 0.3 M KCl, 20 mM MOPS (pH 7.0), 100 $\mu\text{g ml}^{-1}$ bovine serum albumin, and a mixture of protease inhibitors (as for Fig. 1), 2 μM cADPR (except for control), 100 μM EGTA and 21, 74, 98 and 200 μM CaCl_2 , respectively. Nonspecific binding (11% at pCa=5) was determined in the presence of excess (0.5 μM) unlabelled ryanodine. Data represent the average of 6 measurements on 5 to 6 preparations $\pm$ s.e. Asterisk indicates a significant difference at $P < 0.02$ from the control (100%) using *t*-test analyses.

both micromolar *cis* $[\text{Ca}^{2+}]$ and cADPR significantly enhanced the P_o , as compared to that obtained at 10^{-8} M $[\text{Ca}^{2+}]$.

cADPR influenced the cardiac RyRC in a Ca^{2+} -dependent fashion (Fig. 3). At submicromolar extravesicular $[\text{Ca}^{2+}]$, it increased the rate of $^{45}\text{Ca}^{2+}$ efflux (see also Fig. 1*a*). But at free Ca^{2+} above 10^{-7} M , at which $^{45}\text{Ca}^{2+}$ efflux from cardiac vesicles is already activated by extravesicular Ca^{2+} (Ca^{2+} -induced Ca^{2+} release (CICR)), the stimulatory effect by cADPR was diminished (Fig. 3*a*). The Ca^{2+} -dependence of cADPR action on the receptor was also investigated by measuring $[^3\text{H}]$ ryanodine binding (Fig. 3*b*). We found that at non-saturating ligand concentrations, when the amount of bound $[^3\text{H}]$ ryanodine reflects interactions with both the activators and the inhibitors of the channel¹⁹, cADPR significantly augmented ryanodine binding at lower $[\text{Ca}^{2+}]$, where cADPR also activated Ca^{2+} release. On the other hand, at higher extravesicular $[\text{Ca}^{2+}]$ (above 10^{-6} M), when cADPR had no appreciable effect on Ca^{2+} efflux, it did not significantly influence ryanodine binding.

Figure 3*b* also shows that, in contrast to the cardiac SR case, cADPR inhibits $[^3\text{H}]$ ryanodine binding to skeletal SR in the entire pCa range investigated: the reduction in ryanodine binding by cADPR was more pronounced at elevated $[\text{Ca}^{2+}]$ and significant above 10^{-5} M $[\text{Ca}^{2+}]$. The finding that cADPR also influences ryanodine binding in skeletal SR indicates that cADPR does bind to the skeletal receptor channel, but apparently induces a conformational change, which must be different from that in the cardiac isoform as it does not lead to altered Ca^{2+} flux through the skeletal channel.

The results presented here and obtained from Ca^{2+} flux and $[^3\text{H}]$ ryanodine binding studies, as well as from single channel recordings, indicate that cADPR activates the cardiac type of the RyRC and, as a consequence, stimulates Ca^{2+} release from cardiac SR and brain microsomal vesicles. Taken together with the facts that: (1) both cardiac and brain cells produce cADPR¹³ in concentrations²⁰ adequate to activate the RyRC, and (2) there are enzymatic activities to eliminate cADPR (ref. 21; and L.G.M. and J.B., unpublished results), these results strongly suggest that cADPR is a good candidate as an endogenous modulator of the cardiac RyRC isoform in the heart and the brain. Moreover, as both the RyRC⁶⁻¹¹ and cADPR¹³ appear to be common to various other cell types, the modulation of the RyRC by cADPR may have physiological relevance not only in the heart and brain, but in other tissues as well. Activation by cADPR of Ca^{2+} release from intracellular Ca^{2+} stores in pituitary cells²² and sea urchin eggs¹⁴ has been reported. It is possible that cADPR may play a role analogous to that of inositol trisphosphate¹⁵ in Ca^{2+} signalling, but targeting a separate intracellular Ca^{2+} channel, the non-skeletal isoform of the RyRC.

The presence of the RyRC, and the Ca^{2+} responses to

caffeine and ryanodine, have often been linked to possible *in situ* CICR, thought to be the physiological mechanism by which RyRC is activated in cells other than skeletal muscle. But the finding that cADPR activates cardiac RyRC at low submicromolar $[Ca^{2+}]$, when by definition CICR does not operate, suggests that cADPR-dependent activation of the channel might represent a trigger mechanism as an alternative to CICR. □

Received 29 December 1992; accepted 7 April 1993.

- Otsu, K. *et al.* *J. Biol. Chem.* **265**, 13472–13483 (1990).
- Fleischer, S. & Inui, M. *Rev. biophys. Chem.* **18**, 333–364 (1989).
- Lai, F.A. & Meissner, G. *Adv. exp. med. Biol.* **269**, 73–77 (1990).
- Fabiato, A. *Molec. cell. Biochem.* **89**, 135–140 (1989).
- Rios, E. & Pizarro, G. *Physiol. Rev.* **71**, 849–908 (1991).
- Ashley, R.H. *J. Membr. Biol.* **111**, 179–189 (1989).
- Cheek, T.R., Barry, V.A., Berridge, M.J. & Missiaen, L. *Biochem. J.* **275**, 697–701 (1991).
- Fasolato, C. *et al.* *J. Biol. Chem.* **266**, 20159–20167 (1991).
- Mészáros, L.G. & Volpe, P. *Am. J. Physiol.* **261**, C1048–C1054 (1991).
- Shoshan-Barmatz, V., Pressley, T.A., Higham, S. & Kraus-Friedman, N. *Biochem. J.* **276**, 41–46 (1991).
- Giannini, G., Clementi, E., Ceci, R., Marziali, G. & Sorrentino, V. *Science* **257**, 91–94 (1992).

- Lee, H.C., Walseth, T.F., Bratt, G.T., Hayes, R.N. & Clapper, D.L. *J. Biol. Chem.* **264**, 1608–1615 (1989).
- Rusinko, N. & Lee, H.C. *J. Biol. Chem.* **264**, 11725–11731 (1989).
- Gallione, A., Lee, H.C. & Busa, W.B. *Science* **253**, 1143–1146 (1991).
- Berridge, M.J. & Irvine, R.F. *Nature* **341**, 197–205 (1989).
- Ikemoto, N., Ronjat, M. & Mészáros, L.G. *J. Bioenerg. Biomembr.* **21**, 247–266 (1989).
- Mészáros, L.G. & Ikemoto, N. *J. Biol. Chem.* **260**, 16076–16079 (1985).
- Meissner, G.T. *J. Biol. Chem.* **261**, 6300–6306 (1986).
- Chu, A., Diaz-Munoz, M., Hawkes, M.J., Brush, K. & Hamilton, S.L. *Molec. Pharmacol.* **37**, 735–741 (1990).
- Waseth, T.F., Aarhus, R., Zeleznikar, R.J. & Lee, H.C. *Biochem. biophys. Acta* **1094**, 113–120 (1991).
- Gallione, A. *Trends pharmac. Sci.* **13**, 304–306 (1992).
- Koshiyama, H., Lee, H.C. & Tashjian, A.H. *J. Biol. Chem.* **266**, 16985–16988 (1991).
- Colquhoun, D. & Sigworth, F.J. in *Single-channel Recording* (eds Sakmann, B. & Neher, E.) 191–264 (Plenum, New York, 1983).
- Sumida, N., Wang, T., Mandel, F., Froehlich, J.P. & Schwartz, A. *J. Biol. Chem.* **253**, 8772–8777 (1978).
- Edelman, A.M., Hunter, D.D., Hendrickson, A.E. & Krebs, E.G. *J. Neurochem.* **5**, 2609–2617 (1985).
- Fill, M. *et al.* *Biophys. J.* **50**, 471–475 (1990).
- Chu, A. & Stefani, E. *J. Biol. Chem.* **266**, 7699–7705 (1991).
- Hamilton, S.L. *et al.* *Analyt. Biochem.* **183**, 31–41 (1989).
- Rousseau, E., Smith, J.S. & Meissner, G. *Am. J. Physiol.* **253**, C364–C368 (1987).

ACKNOWLEDGEMENTS. We thank H. Lee for cADPR, M. Fill for advice on single channel conductance measurements, and B. Erlich for discussion. This work was supported by grants from the American Heart Association to L.G.M. (Georgia Affiliate) and to A.C. (Texas Affiliate).

Mammalian facilitative hexose transporters mediate the transport of dehydroascorbic acid

Juan Carlos Vera*, Coralía I. Rivas*, Jorge Fischbarg† & David W. Golde*

* Program in Molecular Pharmacology and Therapeutics, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10021, USA

† Departments of Physiology and Cellular Biophysics and Ophthalmology, College of Physicians and Surgeons, Columbia University, New York, New York 10032, USA

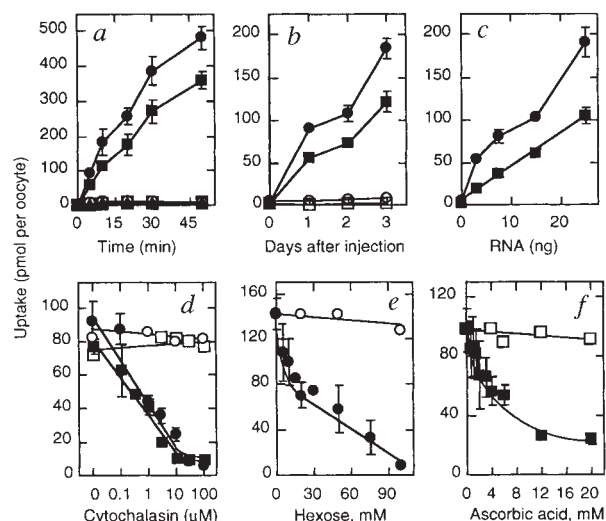
ALTHOUGH vitamin C is critical to human physiology^{1–5}, it is not clear how it is taken up into cells. The kinetics of cell and tissue accumulation of ascorbic acid *in vitro* indicate that the process is mediated by specific transporters at the cell membrane⁶. Some experimental observations have linked the transport of ascorbic acid with hexose transport systems in mammalian cells, although no clear information is available regarding the specific role(s) of these transporters, if any, in this process^{7–16}. Here we use the *Xenopus laevis* oocyte expression system to show that the mammalian facilitative hexose transporters are efficient transporters of the oxidized form of vitamin C (dehydroascorbic acid). Two transport pathways, one with low affinity and one with high affinity for dehydroascorbic acid, were found in oocytes expressing the mammalian transporters, and these oocytes accumulated vitamin C against a concentration gradient when supplied with dehydroascorbic acid. We obtained similar results in experiments using normal human neutrophils. These observations indicate that mammalian facilitative hexose transporters are a physiologically significant pathway for the uptake and accumulation of vitamin C by cells, and suggest a mechanism for the accumulation of ascorbic acid against a concentration gradient.

A major problem associated with studies of ascorbic acid uptake *in vitro* is the instability of the molecule, which is rapidly oxidized to dehydroascorbic acid and further oxidation products¹⁷. We maintained preparations of ascorbic acid at room temperature for several hours in the presence of dithiothreitol without formation of oxidation products (data not shown). The oxidized form of ascorbic acid was produced by

treating samples with ascorbate oxidase in medium containing dithiothreitol, a process followed by measuring the decrease in absorption at 265 nm¹⁸. *Xenopus* oocytes expressing the hexose transport protein isoform glut-1 (refs. 19–20) had markedly increased capacity to accumulate radioactive ascorbic acid when incubated in the presence of freshly prepared dehydroascorbic acid (Fig. 1a). No significant accumulation of radioactivity, however, was observed when the oocytes were incubated with the reduced form of ascorbic acid (Fig. 1a). There was a substantial increase in the uptake of radioactive vitamin as early as 24 h after injecting the RNA (Fig. 1b), an effect that was dose-dependent in oocytes injected with different amounts of RNA (Fig. 1c). The increase in the capacity of the injected oocytes to accumulate the vitamin paralleled the increase in their capacity to accumulate 2-deoxy-D-glucose (Fig. 1). No accumulation of ascorbic acid or deoxyglucose was observed in oocytes injected with an RNA encoding a carboxy-terminal truncated form of the transporter²² which cannot transport glucose. Immunoprecipitation indicated that the corresponding proteins were localized to the plasma membrane (data not shown). Further analysis showed that cytochalasin B (ref. 23), but not cytochalasin E, inhibited the uptake of ascorbic acid and 2-deoxy-D-glucose by these oocytes with $K_i \sim 0.5 \mu\text{M}$ (Fig. 1d). Phloretin²⁴ (100 μM) ablated the accumulation of dehydroascorbic acid by injected oocytes (data not shown). In a related experiment, 2-deoxy-D-glucose, but not L-glucose, inhibited the accumulation of ascorbic acid by the injected oocytes (Fig. 1e) with a $K_i \sim 10 \text{ mM}$. Uptake of the hexose by the expressed transporter was inhibited by dehydroascorbic acid ($K_i \sim 3 \text{ mM}$) (Fig. 1f). Dehydroascorbic acid also exhibited countertransport acceleration²⁴ induced by the presence of a hexose, and preincubation with insulin^{20,21} reversibly enhanced (2-fold) the uptake of dehydroascorbic acid in oocytes expressing the glut-1 protein (data not shown). Oocytes expressing two additional members of the family of mammalian facilitative glucose transporters, the protein isoforms glut-2 and glut-4 (refs. 19–21, 25), also showed a notable increase in their capacity to take up dehydroascorbic acid but not the reduced form of the vitamin (data not shown). Taken together, these results indicate the direct participation of the facilitative glucose transporters in the transport of dehydroascorbic acid by *Xenopus* oocytes expressing the mammalian proteins.

Kinetic studies showed that the uptake of dehydroascorbic acid by oocytes expressing the protein glut-1 was concentration-dependent (Fig. 2a). Lineweaver–Burk analysis revealed an

FIG. 1 Uptake of vitamin C and 2-deoxy-D-glucose in oocytes expressing the glut-1 protein. *a*, Time course of the uptake of reduced ascorbic acid (triangles), dehydroascorbic acid (circles), or 2-deoxy-D-glucose (squares) in oocytes expressing the protein glut-1 (solid symbols), or control oocytes injected with water (open symbols). Uptake assays were carried out 3 d after RNA injection. Data represent the mean $\pm$ s.d. of four groups of 10 oocytes each. *b*, Dependence of uptake on time after RNA injection. Oocytes were injected with RNA encoding the protein glut-1 (solid symbols) or with water (open symbols) and the uptake of dehydroascorbic acid (circles) or 2-deoxy-D-glucose (squares) was measured at different times after the injection, using a 10-min uptake assay. Data represent mean $\pm$ s.d. of four groups of 10 oocytes each. *c*, Dependence on the amount of RNA injected. Oocytes were injected with different amounts of RNA, and the uptake of dehydroascorbic acid (●) or 2-deoxy-D-glucose (■) was measured after 3 d. Data represent the mean $\pm$ s.d. of four to five groups of 10 oocytes each. *d*, Effect of cytochalasins on the uptake of dehydroascorbic acid (circles) or 2-deoxy-D-glucose (squares). Oocytes were assayed for uptake 3 d after injection. Before the uptake assay, oocytes were incubated for 5 min in medium containing the indicated concentrations of cytochalasin B (filled symbols) or cytochalasin E (open symbols). The respective cytochalasins were also present during the 10-min uptake assay. Data represent the mean $\pm$ s.d. of three groups of 10 oocytes each. *e*, Effect of hexoses on the uptake of dehydroascorbic acid. Three days after injection, oocytes were assayed for their capacity to take up dehydroascorbic acid in the presence of the indicated concentrations of 2-deoxy-D-glucose (●) or L-glucose (○). Data represent mean $\pm$ s.d. of four groups of 10 oocytes each. *f*, Effect of ascorbate on the uptake of 2-deoxy-D-glucose. Three days after injection, oocytes were tested for their capacity to take up 2-deoxy-D-glucose in the presence of the indicated concentrations of dehydroascorbic acid (■) or reduced ascorbic acid (□). Data represent the mean $\pm$ s.d. of 4 groups of 10 oocytes each. The source of DNAs, the *in vitro* preparation of RNA, and the collection of oocytes and injections have been described^{20,21}. For uptake assays, oocytes (10 per group) were incubated for up to 50 min in 1 ml modified Barth's saline (MBS) containing 0.1 mM dithiothreitol (DTT)



and 0.1 mM ascorbic acid with 0.5 μ Ci L-[1-¹⁴C]-ascorbic acid (4.74 mCi per mmol; NEN-DuPont). Where indicated, incubation medium also contained 5 units of ascorbate oxidase per ml (Sigma Chemical). Uptake was terminated by washing oocytes with three 7-ml portions of ice-cold MBS containing 0.1 mM phloretin. Oocytes were dissolved in 0.2 ml of 0.5% sodium dodecyl sulphate (SDS), and the incorporated radioactivity assayed by liquid scintillation counting. The standard uptake assay comprised a 10-min incubation period. Uptake of 2-deoxy-D-glucose was assayed as described previously, using 0.5 mM 2-deoxy-D-glucose and 5 μ Ci ml⁻¹ of 2-[1,2-³H (N)]-deoxy-D-glucose (26.2 Ci per mmol, NEN-DuPont) and a 10-min incubation period^{20,21}.

apparent K_m of 3.5 mM and a V_{max} of 190 pmol per oocyte per min. The presence of a second, high-affinity saturable component was apparent when oocytes were incubated at concentrations of dehydroascorbic acid of less than 300 μ M. After correcting for the contribution of the low-affinity component (Fig. 2b), Lineweaver-Burk analysis revealed an apparent K_m for dehydroascorbic acid of 60 μ M, with a corresponding V_{max} of

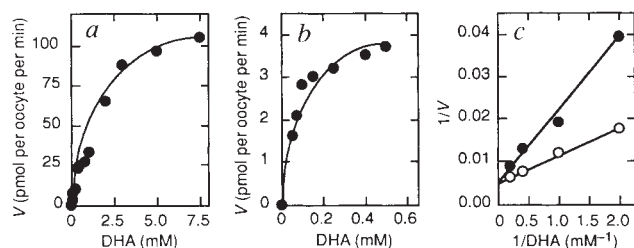


FIG. 2 Kinetic analysis of dehydroascorbic acid (DHA) uptake by oocytes expressing the glut-1 protein. *a*, Concentration dependence. Uptake was measured in the presence of different concentrations of dehydroascorbic acid in a 10-min assay. *b*, Saturation curve for transport of substrate by the high-affinity pathway. Oocytes were incubated in the presence of concentrations of dehydroascorbic acid ranging from 50 μ M to 1 mM and a curve of uptake versus initial concentration of dehydroascorbic acid was drawn. A straight line was constructed from the origin to the value of uptake observed at a concentration of dehydroascorbic acid of 1 mM. Single-point values on this line were subtracted point by point from the corresponding values on the curve of uptake versus concentration of ascorbic acid. *c*, Effect of 2-deoxy-D-glucose on the kinetic parameters of the uptake of dehydroascorbic acid. Oocytes were incubated in MBS containing different concentrations of ascorbic acid ranging from 0.5 to 5 mM, in the presence (●) or absence (○) of 10 mM 2-deoxy-D-glucose.

4.8 pmol per oocyte per min. 2-Deoxy-D-glucose inhibited the transport of 1 mM (Fig. 1e) and 10 μ M (data not shown) dehydroascorbic acid with a $K_i \sim 10$ mM. We also found that 2-deoxyglucose competitively inhibited the transport of dehydroascorbic acid mediated by the low- (Fig. 2c) and the high-affinity sites (data not shown), with a K_i of 12 mM. Conversely, no significant effect on oocyte accumulation of 2-deoxy-D-glucose by the oocytes was seen at micromolar concentrations of dehydroascorbic acid (Fig. 2b, and data not shown). But at concentrations above 0.5 mM, dehydroascorbic acid competitively inhibited the uptake of 2-deoxy-D-glucose, with an apparent K_i of 4 mM. Others^{16,26} have described two transport activities with different affinities for ascorbic acid in human neutrophils, which was thought to suggest the existence of two different transporters in those cells. Our findings instead indicate that such observations may be due to the presence of a single transport protein with two functional sites that have different affinities for the transported substrate. It is difficult to estimate the relative contributions of each site to the net transport of the vitamin under physiological conditions, but the low concentrations of ascorbic acid present in blood^{6,27} favour an important role for the high-affinity site and suggests that the transporters are not saturated.

The observation that some cells accumulate high concentrations of ascorbic acid against a concentration gradient has been used as an argument against the participation of facilitative glucose transporters in the cellular uptake of ascorbic acid^{16,26,28,29}. We found that oocytes expressing glut-1 accumulated about 2.5 times more radioactive vitamin than the amount expected from simple uptake into a cell with an average internal volume of about 0.4 μ l (for a full-grown stage-6 oocyte) (Fig. 1a). This effect was dependent on the concentration of dehydroascorbic acid during the incubation, and was most noticeable in the low micromolar concentration range (Fig. 3a). Subcellular

FIG. 3 Accumulation of ascorbic acid in oocytes expressing the glut-1 protein. *a*, Oocytes were incubated for 60 min in MBS containing different concentrations of radioactive dehydroascorbic acid, washed, and the radioactivity in the oocytes determined by scintillation counting. An internal volume of 0.4 μ l per oocyte was used to express the accumulated radioactivity against a concentration gradient. Data are expressed as the ratio of ascorbic acid accumulated in the oocytes to the concentration in the extracellular medium. The dashed line indicates a ratio equal to one. A ratio higher than one indicates accumulation of radioactivity against a concentration gradient. Data represent measurements obtained from groups of 10 oocytes each. *b*, Oocytes were incubated for 60 min in MBS containing 0.1 mM radioactive dehydroascorbic acid. At the end of the incubation period, oocytes were rapidly washed in radioactivity-free medium and crushed by centrifugation in a microfuge at 10,000*g* for 5 min. The pellet and supernatant were recovered and the radioactivity associated with each fraction was measured by scintillation spectrometry. Data represent means of three groups of 20 oocytes each. *c*, Oocytes were incubated for 60 min in the presence of 0.1 mM radioactive ascorbic acid and divided into two groups. One group was injected with 50 nl MBS containing 50 units μ l⁻¹ of ascorbate oxidase (○). The second group was injected with 50 nl MBS containing 0.1 mg ml⁻¹ bovine serum albumin (●). Oocytes were transferred to vitamin-free medium and radioactivity remaining in the oocytes measured for the times indicated. Data are presented as per cent of the radioactivity initially present in the oocytes at the beginning of the final incubation period, and represent the mean of three groups of 10 oocytes each.

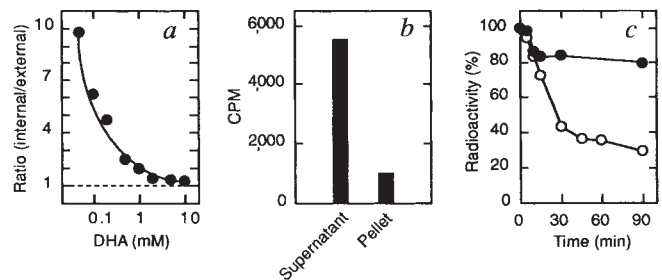
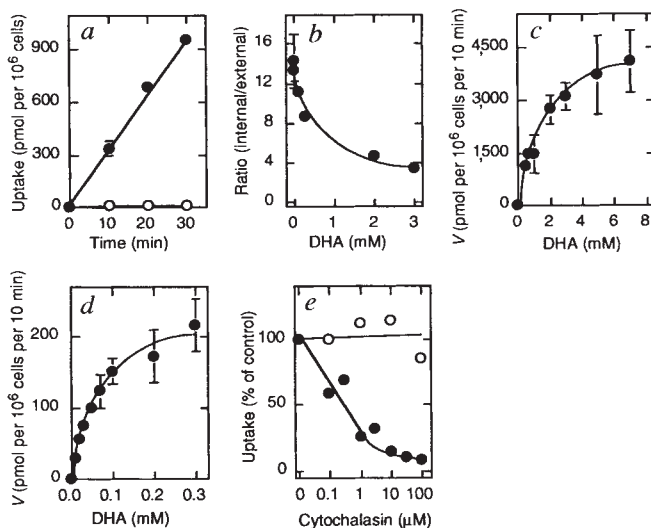


FIG. 4 Uptake of dehydroascorbic acid by normal human neutrophils. *a*, Time-course of the uptake of reduced (○), or oxidized (●) ascorbic acid. Results correspond to the mean $\pm$ s.d. of four samples. *b*, Accumulation of ascorbic acid as function of concentration. Cells were incubated for 10 min in medium containing increasing concentrations of dehydroascorbic acid, and accumulated radioactivity was determined by scintillation spectrometry. An estimated value of 0.30 μ l per 10⁶ cells was used to express the accumulated radioactivity on a concentration basis. Data are expressed as the ratio of ascorbate accumulated in the neutrophils to the concentration in the extracellular medium, and represent the mean $\pm$ s.d. of four samples. *c*, Substrate saturation curve for the transport of dehydroascorbic acid by the low-affinity pathway. Uptake was measured in a 10-min assay for concentrations of dehydroascorbic acid ranging from 0.5 to 7 mM. Data represent means $\pm$ s.d. of four samples. *d*, Substrate saturation curve for the transport of dehydroascorbic acid by the high-affinity pathway. Uptake was measured for concentrations of dehydroascorbic acid ranging from 0.01 to 1 mM and data were corrected for the contribution of the low-affinity pathway as indicated in Fig. 2. Data represent the mean $\pm$ s.d. of four samples. *e*, Effect of cytochalasins on the uptake of dehydroascorbic acid. Cells were incubated for 10 min in medium containing the indicated concentrations of cytochalasin B (●) or cytochalasin E (○) before the 10-min uptake assay. The cytochalasins were also present during the uptake assay. Data represent the mean of duplicate samples. Leukopacks were provided by the New York Blood Center and neutrophils were purified by Ficoll-Hypaque density centrifugation. Cells were suspended in incubation medium at a final concentration of 2×10^6 cells ml⁻¹. Cell viability, as determined by exclusion of trypan blue, was $> 95\%$. For uptake assays, 0.5 ml cells was added to 0.5 ml incubation medium containing 0.1 mM DTT. When appropriate, the medium also contained ascorbic acid, L-[¹⁴C]ascorbic acid (0.1–0.5 μ Ci), ascorbate oxidase (5 units ml⁻¹), cytochalasin B, cytochalasin E, D-glucose, or L-glucose. In the standard assay, the cell mixture was incubated for 10 min at room temperature. Cells were collected by centrifugation and washed twice in cold (4°C) stopping solution (PBS without Ca²⁺ and Mg²⁺, containing 100 μ M phloretin and 20 μ M cytochalasin B). Cellular pellets were solubilized in 20 mM Tris-HCl, pH 8.0, containing 1% SDS, and cell-associated radioactivity determined by scintillation spectrometry. A sample in which the cells were immediately centrifuged in the presence of stopping solution was used as a control for nonspecifically trapped radioactivity in the cell pellet. For determination of cell volume, neutrophils were suspended in medium supplemented with 0.1 mM 3-O-methyl-D-glucose and traces of 3-O[methyl-³H]-D-glucose, incubated for 0 or 4 h at room temperature and processed as described.



fractionation studies indicated that there was no nonspecific binding of the vitamin to the abundant yolk protein present in oocytes (Fig. 3b). We postulated that the excess accumulation of ascorbate was a result of intracellular modification into a form that is not transported out by the facilitative glucose transporter. To test this hypothesis, oocytes preloaded with the radioactive vitamin were injected with ascorbic acid oxidase, after which we measured the efflux of radioactive material. Control oocytes were injected with bovine serum albumin. The injection of ascorbate oxidase induced a rapid efflux of radioactive material that could not be explained by nonspecific leakage resulting from the injection procedure (Fig. 3c). Preloaded oocytes were also subjected to methanol extraction and the extracts analysed for ascorbate by thin-layer chromatography. No significant amount of dehydroascorbic acid was observed in oocytes preloaded with the vitamin, and the radioactivity accumulated in the oocytes migrated mainly in the expected position for the reduced form of ascorbic acid (data not shown). Thus, ascorbic acid can be accumulated against a concentration gradient by a mechanism involving facilitated uptake of the oxidized form and retention of the reduced ascorbic acid.

We also analysed the transport of ascorbic acid by human neutrophils^{9,16,26,29–31}, and the results closely matched those from *Xenopus* oocytes expressing mammalian glucose transporters. Under conditions that prevented spontaneous oxidation of ascorbic acid, the neutrophils did not take up the reduced form of ascorbic acid (Fig. 4a), but rapidly transported dehydroascorbic acid (Fig. 4a). Neutrophils accumulated ascorbic acid against a concentration gradient when exposed to dehydroascorbic acid (Fig. 4b) and two transport activities were detectable with different affinities for dehydroascorbic acid (Fig. 4c–d). The low-affinity component showed an apparent K_m of 2.5 mM, and the high-affinity component showed a corrected K_m of 60 μ M. Cytochalasin B completely inhibited the transport of dehydroascorbic acid with a K_i of 0.6 μ M (Fig. 4e), and cytochalasin B, but not cytochalasin E, inhibited the transport of 2-deoxy-D-glucose with a K_i of 0.5 μ M (data not shown). Finally, we documented competitive inhibition of the transport of ascorbic acid by glucose with a K_i of 9 mM (data not shown). These observations strongly support the concept that the glucose transporters are a physiologically significant pathway for the uptake of ascorbate in human cells.

In conclusion, our results with *Xenopus* oocytes and human neutrophils provide evidence indicating a direct role for the mammalian facilitative glucose transporters in the transport of ascorbic acid. Our data are consistent with the hypothesis that vitamin C is transported by the glucose transporters as dehydroascorbic acid and then is rapidly reduced to ascorbic acid in the interior of cell, a mechanism that allows for the accumulation of ascorbic acid against a concentration gradient. □

Received 20 January; accepted 13 April 1993.

- Crandon, J. H. & Lund, C. C. *New Engl. J. Med.* **223**, 353–369 (1940).
- Hodges, R. E., Baker, E. M., Hood, J. H., Sauberlich, H. E. & Baker, E. M. *Am. J. clin. Nutr.* **22**, 535–548 (1969).
- Hodges, R. E., Baker, E. M., Hod, J. H., Sauberlich, H. E. & Baker, E. M. *Am. J. clin. Nutr.* **24**, 432–443 (1971).
- Englard, S. & Steifter, S. A. *Rev. Nutr.* **6**, 365–406 (1986).
- Padh, H. *Biochem. Cell Biol.* **68**, 1166–1173 (1990).
- Rose, R. C. *Biochim. biophys. Acta* **947**, 335–366 (1988).
- Siliprandi, L., Vanni, P., Kessler, M. & Semenza, G. *Biochim. biophys. Acta* **552**, 129–142 (1979).
- Toggenburger, G. *et al. Biochim. biophys. Acta* **646**, 433–443 (1981).
- Bigley, R., Wirth, M., Layman, D., Riddle, M. & Stankova, L. *Diabetes* **32**, 545–548 (1983).
- Bianchi, J., Wilson, F. A. & Rose, R. C. *Am. J. Physiol.* **250**, G461–G468 (1986).
- Ingermann, R. L., Stankova, L. & Bigley, R. H. *Am. J. Physiol.* **250**, C637–C641 (1986).
- Rose, R. C. *Am. J. Physiol.* **250**, F627–F632 (1986).
- Padh, H. & Aleo, J. J. *Biochim. biophys. Acta* **901**, 283–290 (1987).
- McLennan, S. *et al. Diabetes* **37**, 359–361 (1988).
- Wilson, J. X. & Dixon, J. S. *J. Membr. Biol.* **111**, 83–91 (1989).
- Washko, P. & Levine, M. *J. biol. Chem.* **267**, 23568–23574 (1992).
- Winkler, B. S. *Biochim. biophys. Acta* **925**, 258–264 (1987).
- Wunderling, M., Paul, H. H. & Lohmann, W. *Biol. Chem. Hoppe-Seyler* **367**, 1047–1054 (1986).

- Birnbaum, M., Haspel, H. & Rosen, O. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5784–5788 (1986).
- Vera, J. C. & Rosen, O. *Molec. cell. Biol.* **9**, 4187–4195 (1989).
- Vera, J. C. & Rosen, O. *Molec. cell. Biol.* **10**, 743–751 (1990).
- Oka, Y. *et al. Nature* **345**, 550–553 (1990).
- Jung, C. Y. & Rampal, A. L. *J. biol. Chem.* **252**, 5456–5463 (1977).
- Wheeler, T. J. & Hinkle, P. C. *A. Rev. Biochem.* **47**, 503–517 (1985).
- Thorens, B., Sarkar, H., Kaback, H. & Lodish, H. *Cell* **55**, 281–290 (1988).
- Washko, P., Rotrosen, D. & Levine, M. *J. biol. Chem.* **264**, 18996–19002 (1989).
- Dhariwal, K. R., Hartzell, W. O. & Levine, M. *Am. J. clin. Nutr.* **54**, 712–716 (1991).
- Moser, U. & Weber, F. *Int. J. Vit. Nutr. Res.* **54**, 47–53 (1984).
- Bergsten, P. *et al. J. biol. Chem.* **265**, 2584–2587 (1990).
- Bigley, R. H. & Stankova, L. *J. exp. Med.* **139**, 1084–1092 (1974).
- Raghoobar, M., Huisman, J. A. M., van den Berg, W. B. & van Ginnekan, C. A. M. *Life Sci.* **40**, 499–510 (1987).

ACKNOWLEDGEMENTS. This work was supported by grants from the NIH, the Memorial Sloan-Kettering Cancer Center institutional funds and the Schultz Foundation.

Inhibition of protein kinase C by alcohols and anaesthetics

Simon J. Slater, Kingsley J. A. Cox, Joseph V. Lombardi, Cojen Ho, Mary Beth Kelly, Emanuel Rubin & Christopher D. Stubbs*

Department of Pathology and Cell Biology, Thomas Jefferson University, Philadelphia, Pennsylvania 19107, USA

DESPITE almost a century of research, the mechanism of anaesthesia remains obscure and there is still no agreement on the location of the site(s) of action^{1–7}. Because the potencies of general anaesthetics increase in proportion to their solubility in olive oil, this led to a consensus that the site is within the cell membrane^{8–10}. This led to theories that lipid bilayer perturbation was the primary event, which was then transmitted to a membrane protein¹¹. But at the concentrations used clinically, such perturbations are small³. A plausible site would be in or on ion channels at the synapse, where a number of modulatory effects have been described⁶. A possible location for such a site would be at the protein–lipid interface^{5,12,13}. We report here that anaesthetics inhibit protein kinase C, a key component in signal transduction. The potency is a linear function of the octanol–water partition coefficient (the Meyer–Overton rule of anaesthesia). The effect was obtained in a lipid-free assay, implicating a hydrophobic site in the protein, supporting the contention that a (membrane) protein may be a target for anaesthetic interactions^{14–17}. In a lipid-dependent assay, a potential role of lipids in the protein-site model was demonstrated. The inhibition was absent in the isolated catalytic domain, suggesting that the site of inhibition is on the regulatory subunit, which is unique to protein kinase C.

Protein kinase C (PKC) regulates synaptic function by phosphorylation of membrane proteins, including ion channels, and its modulatory role in neuronal signal transduction is well documented¹⁸. PKC from rat brain, activated in a lipid-free assay by protamine sulphate¹⁹, was inhibited by members of the homologous series of aliphatic *n*-alkanols (Fig. 1). In agreement with the Meyer–Overton rule, the concentration required for a 50% inhibition of the enzyme (IC_{50}) increased as a linear function of the alcohol chain length, at least up to decanol, and was also a linear function of the octanol–water partition coefficient (Fig. 1, inset). Further, the value of the free energy for the transfer of an *n*-alcohol from the buffer to the binding site was -700 cal per mol per methylene unit (determined from the slope of a plot of chain length against $-RT \ln IC_{50}$, where R is the gas constant and T is temperature), within the range for the transfer of a methylene unit from water to a non-aqueous phase¹. Thus, although some effects of anaesthetics on PKC

*To whom correspondence should be addressed.

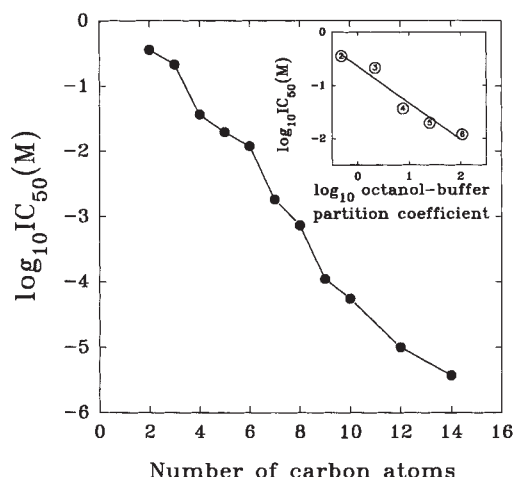


FIG. 1 The dependence of the potency of PKC inhibition (IC_{50}) by a homologous series of *n*-alkanols in the presence of the PKC activator protamine sulphate. IC_{50} values were calculated from PKC activities determined over a range of alcohol concentrations. Inset: potency of inhibition by alcohols as a function of the buffer-octanol partition coefficients²⁸.

METHODS. The activity of PKC, purified from rat brain²⁹ or obtained from Lipidex (Westfield, NJ; purified using the same procedure), was determined as the incorporation of ^{32}P from [γ - ^{32}P]ATP (New England Nuclear) into histone type III-S (Sigma), as previously described²⁹, except that the reaction mixture included 0.4 mg ml^{-1} protamine sulphate (Sigma) in place of lipids. Long-chain alcohols and volatile anaesthetics were added from intermediate (sealed) suspensions in buffer, and assays were done in sealed tubes with limited air space, by which means losses were found to be negligible.

have been previously reported²⁰⁻²², the results provide crucial evidence that alcohols are acting at a hydrophobic site on the enzyme.

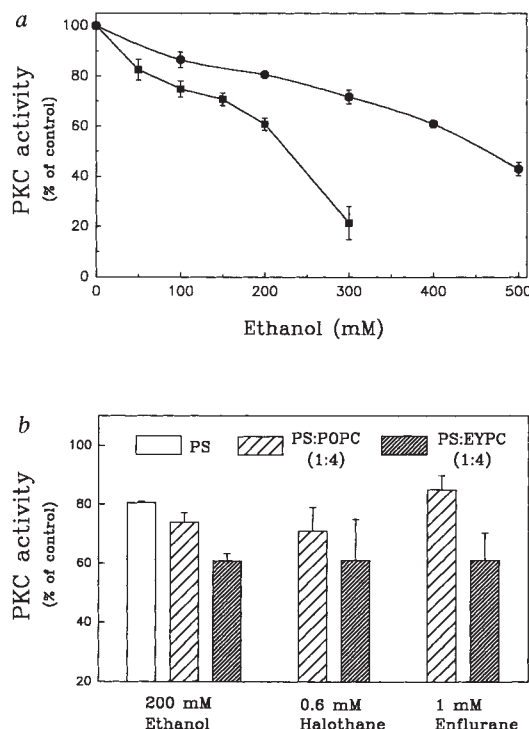
It is common to assess the relevance of an anaesthetic effect by comparing the IC_{50} values with the amount of anaesthetic that is 50% effective in bringing about anaesthesia (ED_{50}). The IC_{50} values for the inhibition of PKC, activated by protamine sulphate, for enflurane and halothane, were 0.95 and 0.8 mM, respectively. The latter value can be compared with its ED_{50} of 0.23 mM for the tadpole righting reflex²³. Although the IC_{50} values in Fig. 1 are above those required for anaesthesia, this experiment was intended to show that the inhibition could be obtained without lipids. To investigate potency it was necessary to use lipids to activate the PKC to mimic the natural situation better.

Using lipid vesicles of brain-phosphatidylserine (brain-PS) to activate PKC, it was confirmed that the inhibition again followed the Meyer-Overton rule (results not shown). This was done only up to hexanol because for longer-chain alcohols lipid

vesicles can seriously deplete the free alcohol concentration. To study lipid dependency, brain-PS/POPC (monounsaturated phosphatidylcholine) (1:4 molar) vesicles and brain-PS/egg-PC (polyunsaturated phosphatidylcholine) (1:4) vesicles, where the PS level corresponded to that in synaptic membranes, were compared with 100% brain-PS vesicles. The PKC activities were 28.8, 10.8 and $44.1 \text{ pmol ATP min}^{-1} \mu\text{g}^{-1}$ enzyme, respectively. The potencies of inhibition obtained for ethanol were in the order: brain-PS/egg-PC > brain-PS/POPC > brain-PS (Fig. 2). For brain-PS/egg-PC vesicles, the IC_{50} for ethanol was 250 mM, a value that compares with 190 mM for the tadpole ED_{50} (ref. 23). An increased potency of inhibition with egg-PC as compared to POPC was also seen with the volatile anaesthetics, halothane and enflurane (Fig. 2b). Such comparisons between the potency of inhibition of PKC in model vesicle systems and the anaesthetic potency on the tadpole are far from ideal. Nevertheless, the PKC inhibition falls within a concentration range compatible with clinically relevant anaesthetic concentrations and the results also show that the anaesthetic site on PKC is

FIG. 2 The inhibitory effects of anaesthetics were dependent on lipid composition. *a*, Greater inhibition by ethanol was obtained with egg-PC:brain-PS (4:1, molar) vesicles (squares), compared to brain-PS alone (circles). *b*, Comparison of the inhibition obtained with ethanol, halothane and enflurane (results are mean of triplicate determinations $\pm$ s.d.).

METHODS. Activity was determined with lipid vesicles ($150 \mu\text{M}$) prepared as previously described²⁵. The assay was otherwise as described in the legend to Fig. 1 (without protamine sulphate) and with $0.6 \mu\text{M}$ TPA (Sigma) to activate the PKC.



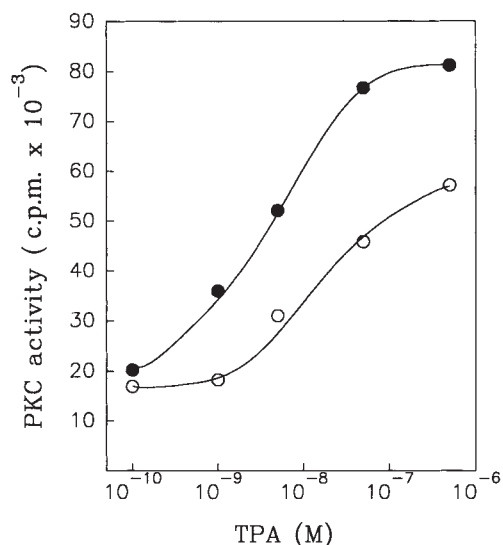


FIG. 3 The dependence of PKC inhibition on TPA concentration, expressed as the ^{32}P , incorporated into histone under assay conditions as described in the legend to Fig. 1, for brain-PS vesicles, with (empty circles) and without (filled circles) 50 mM butanol.

modulated by lipids, although the critical features remain to be explored.

To investigate whether the inhibition of PKC in lipid vesicles was due to the well-known lipid bilayer fluidizing effect of anaesthetics²⁴, a comparison was made between vesicles with two widely differing membrane fluidities, a mixture of dioleoylphosphatidylcholine (DOPC) and brain-PS (4:1, molar) and POPC and brain-PS (4:1), respectively. Using diphenylhexatriene²⁵ to assess fluidity, the former vesicles were found to be more fluid (fluorescence anisotropies were 0.103 ± 0.001 and 0.115 ± 0.001 , respectively), yet the activity of PKC was fourfold greater. This, and the fact that the inhibition was found in the absence of lipids, argues against effects on fluidity being responsible for the inhibition.

The major regions of PKC that might contain a site at which anaesthetics interact, include the regulatory domain containing the diacylglycerol (DAG) binding site, the protein-lipid interface, and the catalytic domain. After insertion into the membrane, there is a further conformational change induced by DAG binding, resulting in activation. The results in Fig. 3 show that the potency of inhibition by alcohols is dependent on the concentration of the phorbol ester 12-*O*-tetradecanoylphorbol-13-acetate (TPA), a commonly used analogue of DAG, which also activates PKC. If the binding site had been exposed in the inactive (cytosolic) form of PKC, the potency of inhibition

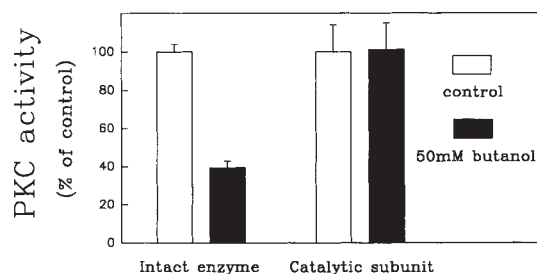


FIG. 4 The inhibition of the catalytic subunit of PKC, prepared as previously described²⁶, was not inhibited by butanol. The activity of the intact enzyme was determined as described in the legend to Fig. 1 in the presence of protamine sulphate. The activity of the catalytic subunit was determined without protamine sulphate or lipids, neither being required for activity²⁶ (results are mean of triplicate determinations $\pm$ s.d.).

would have been independent of the TPA concentration. Therefore, the result suggests that the anaesthetic binding site is revealed only as the enzyme undergoes the conformational change leading to activation. By sharp contrast to the situation with intact PKC, when the catalytic domain separated by proteolytic cleavage was examined, the activity, now freed from lipid or other cofactor requirements²⁶, was unaffected by butanol, even at the high level of 50 mM (Fig. 4). These findings suggest that the site of anaesthetic interaction is located on the regulatory domain. Because among kinases in general it is the catalytic domain that is conserved to a high degree²⁷, this suggests that the inhibition may be unique to PKC. This does not preclude anaesthetics influencing other kinases; however, distinct mechanisms might be anticipated.

Although a direct inhibitory effect on PKC itself was found in this study, demonstration of a direct effect using intact neurons or even cell membrane preparations may not be trivial²². For instance, PKC activation requires the release of intracellular Ca^{2+} and production of DAG, both complex processes that could be influenced by anaesthetics distal to PKC itself. Therefore, to understand the mechanism of direct effects on PKC, it may be preferable to study the isolated enzyme with model vesicle systems.

These data support the protein site hypothesis, but also indicate a potential role for membrane lipids in modulating the anaesthetic effect, possibly by influencing the conformation of the form of PKC inserted into the lipid bilayer. If the binding site is exposed to the bilayer interior it is also possible that the partitioning of the anaesthetic into the lipid bilayer may play a role in regulating its interaction with PKC. The finding of a protein site on PKC suggests that this, and other such sites, may be important in the mechanism of anaesthesia. □

Received 30 November 1992; accepted 13 April 1993.

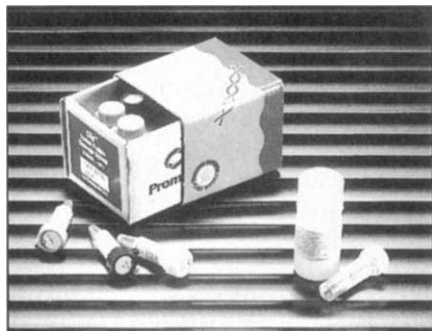
- Seeman, P. *Pharmac. Rev.* **24**, 583–639 (1972).
- Haydon, D. A., Hendry, B. M., Levinson, S. R. & Requena, J. *Nature* **268**, 356–358 (1977).
- Franks, N. P. & Lieb, W. R. *Nature* **300**, 487–493 (1982).
- Dluzewski, A. R., Halsey, M. J. & Simmonds, A. C. *Molec. Aspects Med.* **6**, 459–573 (1983).
- Miller, K. W. *Int. Rev. Neurobiol.* **27**, 1–61 (1985).
- Rubin, E., Miller, K. W. & Roth, S. H. *Ann. N.Y. Acad. Sci.* **625** (1991).
- Matthews, R. *Science* **255**, 156–157 (1992).
- Meyer, H. *Arch. exp. Path. Pharmac.* **42**, 109–118 (1889).
- Meyer, H. *Arch. exp. Path. Pharmac.* **46**, 338–346 (1901).
- Overton, E. *Studien über die Narkose* (Fischer, Jena, 1901).
- Richards, C. D. *et al. Nature* **276**, 775–779 (1978).
- Fraser, D. M., Louro, S. R. W., Horvath, L. I., Miller, K. W. & Watts, A. *Biochemistry* **29**, 2664–2669 (1990).
- Wood, S. C., Forman, S. A. & Miller, K. W. *Molec. Pharmac.* **39**, 332–338 (1991).
- Franks, N. P. & Lieb, W. R. *Nature* **310**, 599–601 (1984).
- Franks, N. P. & Lieb, W. R. *Nature* **316**, 349–351 (1985).
- Curry, S., Leib, W. R. & Franks, N. P. *Biochemistry* **29**, 4641–4652 (1990).
- Moss, G. W. J., Franks, N. P. & Leib, W. R. *Proc. natn. Acad. Sci. U.S.A.* **88**, 134–138 (1990).
- Bell, R. M. & Burns, D. J. *J. biol. Chem.* **266**, 4661–4664 (1991).
- Inoue, M., Kishimoto, A., Takai, Y. & Nishizuka, Y. *J. biol. Chem.* **252**, 7610–7616 (1977).
- Mikawa, K. *et al. J. int. Med. Res.* **18**, 153–160 (1990).
- Trudell, J. R., Costa, A. K. & Hubbell, W. L. *Ann. N.Y. Acad. Sci.* **625**, 743–746 (1991).
- Tas, P. W. & Koschel, K. *Biochim. biophys. Acta* **1091**, 401–404 (1991).
- Alfimoff, J. K., Firestone, L. L. & Miller, K. W. *Br. J. Pharmac.* **96**, 9–16 (1989).
- Trudell, J. R., Hubbell, W. L. & Cohen, E. N. *Biochim. biophys. Acta* **291**, 321–327 (1973).
- Stubbs, C. D., Kinoshita, K., Munkonge, F. M., Quinn, P. J. & Ikegami, A. *Biochim. biophys. Acta* **775**, 374–380 (1984).
- Lee, M. & Bell, R. M. *J. biol. Chem.* **261**, 14867–14870 (1986).
- Hanks, S. K., Quinn, A. M. & Hunter, T. *Science* **241**, 42–52 (1992).
- Leo, A., Hansch, C. & Elkins, D. *Chem. Rev.* **71**, 525–616 (1971).
- Kitano, T., Go, M., Kikkawa, U. & Nishizuka, Y. *Meth. Enzym.* **124**, 349–352 (1986).

ACKNOWLEDGEMENTS. This work was supported by grants from the National Institute of Alcohol Abuse and Alcoholism.

Reagent round-up

Featured this week — a fusion protein cleavage system, an enzyme immunoassay kit for the quantitative determination of 12(S)-hydroxyeicosatetraenoic acid levels and a new mRNA detection system.

THE Clip fusion protein cleavage system from Promega provides three proteases with different cleavage specificities, a positive control protein and a protocol to assist in the development of a strategy to cleave a fusion protein from its purification tag under non-denaturing conditions (*Reader Service No.*



Promega's Clip fusion protein cleavage system.

100). The three proteases included are Factor Xa, trypsin and clostripain. Promega says that Factor Xa has a specific recognition site built into a number of expression vectors (including PinPoint Xa Vectors), but for various reasons the site may not always be accessible. Trypsin and clostripain would not conventionally be tried to cleave fusion proteins but can be successfully used under controlled conditions. The system is compatible with cleaving proteins produced using the company's PinPoint Xa protein purification system. The Clip fusion protein cleavage system provides sufficient reagents for the cleavage of up to 1 mg of fusion protein with each protease.

Packard announces the introduction of its new Ultima-Flo AP cocktail for flow scintillation analysis (*Reader Service No.* 101). This biodegradable cocktail accepts



Ultima-Flo AP — Packard's newest scintillation cocktail.

up to 2.0 M ammonium phosphate gradients at cocktail-to-sample ratios as low as 2:1. Because of its low viscosity, Packard says that Ultima-Flo mixes rapidly with sample, and eliminates pressure spikes. In addition, the company claims that Ultima-Flo has excellent quench resistance, and will not gel even at 1:1 ratios. Its low vapour pressure and high flashpoint (117°C) ensure its safety for shipping, storage and laboratory use without the need for a fume hood. Ultima-Flo is also suitable for up to 1:1 ratios of 50 per cent acetonitrile, 50 per cent methanol, water, 0.25 M PBS, 0.1 M NaOH and 1 M guanidine. It will accept higher molar concentrations at cocktail-to-sample ratios ranging from 2:1 to 5:1.

Assorted assays

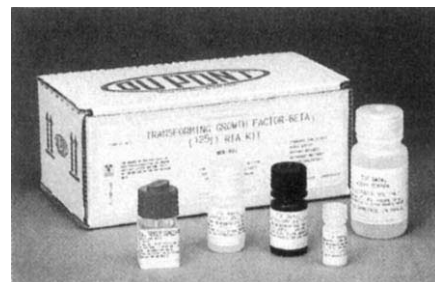
A 12(S)-HETE competitive enzyme immunoassay kit for the quantitative determination of 12(S)-hydroxyeicosatetraenoic acid levels in extracted biological fluids is now available from Advanced Magnetics (*Reader Service No.* 102).



AMI's 12-HETE enzyme immunoassay kit.

This compound has been shown to be chemotactic and chemokinetic for polymorphonuclear leukocytes and vascular smooth muscle cells. 12(S)-HETE acts as a second messenger in angiotensin-II induced aldosterone production. Precision-coated wells supplied with each kit are designed to provide the user with consistency and flexibility. Advanced Magnetics says that this enzyme immuno-assay is optimized for sensitivity with a detection of 0.17 ng ml⁻¹. Assay results can be achieved within 24 h by an overnight incubation. The kit contains sufficient reagents for 96 tests, including a standard curve.

The new radioimmunoassay kit from Du Pont's NEN Research Products measures activated transforming growth factor-beta 1 (TGF-B1) in conditioned tissue culture media and other biological fluids (*Reader Service No.* 103). TGF-B1 belongs to a fam-



Du Pont's RIA measures transforming growth factor.

ily of closely related growth regulatory proteins that serve multifunctional roles, and has been implicated in embryogenesis, tissue repair, inflammation, angiogenesis, immunosuppression and the pathogenesis of certain fibrotic diseases. Du Pont says that it also plays an important role in mechanisms involving wound healing. Studies show that it can up-regulate the synthesis and deposition of extracellular matrix proteins, while down-regulating proteases for degradation of extracellular matrix components. The development by Du Pont of a high affinity antibody to TGF-B1 and high specific activity iodinated TGF-B1 tracer provides researchers with a reproducible radioimmunoassay that is designed to take only six hours to complete. The assay measures TGF-B1 levels from 12.5 to 200 pM (31.25 pg per 100 µl to 500 pg per 100 µl) with percentage B50 at 60 pM (150 pg per 100 µl). Included are all the reagents that are needed to perform the assay, including an I-125-labelled TGF-B1 tracer, a high affinity primary TGF-B1 antibody, a standard, secondary antibody and assay buffer for 100 assays.

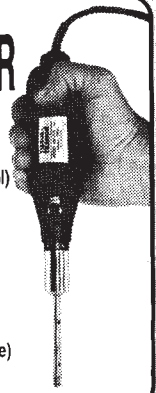
Boehringer Mannheim has recently introduced an ELISA to estimate the soluble CD4 (sCD4) molecule in serum, plasma, cell culture supernatants or cell lysates (*Reader Service No.* 104). The measurement of sCD4 concentrations should prove to be useful in the study of immune regulatory functions and disfunctions. The assay is based on the 'sandwich' principle using two mouse monoclonal antibodies that are directed against two different epitopes of the CD4 glycoprotein. During the first incubation step, sCD4 from standards and samples is bound simultaneously by the antibody-biotin conjugate and the antibody-HRP conjugate (one-step system). This complex binds tightly to the streptavidin-coated microtitre plate. After rinsing the wells to remove unbound compounds, ABTS substrate solution is added and the intensity of the resulting col-

ADVERTISEMENTS

Plant and Animal Tissue

HOMOGENIZER and EMULSIFIER

- Disrupts and disperses small samples (1/2-100 ml)
- Both mechanical and ultrasonic action
- 5,000-30,000 rpm
- Built-in speed control
- Easy to clean
- Best price (\$485 complete)

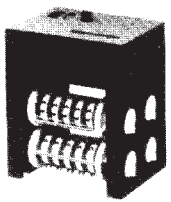


The **TISSUE-TEAROR™**

Manufactured By:
BIOSPEC PRODUCTS • PH/FAX (918) 336-3363
P.O. BOX 722 • BARTLESVILLE, OKLAHOMA 74005

Reader Service No. 15

DIABERM™ EQUILIBRIUM DIALYSIS SYSTEM



The gentlest method to quantify the binding parameters and thermodynamic date between ligands and biopolymers. Results come closest to the actual values. Sample volume from 0.2 ml to 5.0 ml. Up to 20 experiments in parallel.

DIABERM

Carste
P.O. Box 650128
D-5000 Munich 65,
FRG.

Reader Service No.45

1994 KEYSTONE SYMPOSIA on Molecular and Cellular Biology

ATTENTION

The Keystone Symposia is making their 1993 Abstract Volumes available to the scientific community. The cost is \$20.00 plus postage (U.S. dollars) per volume.

For more information, please contact:

Richard D. Handy
Promotion Manager
Keystone Symposia
Drawer 1630
Silverthorne, CO 80498

Phone: (303) 262-1230 • Fax: (303) 262-0311

Reader Service No.11

our development is measured photo-metrically. Colour intensity corresponds to the amount of bound CD4 from standards and samples. Boehringer Mannheim says that the sensitivity of the assay is $\geq 0.25 \text{ ng ml}^{-1}$. All the required reagents and a microtitre plate precoated with streptavidin are included in the kit, which is sufficient for 96 tests.

Immunological tools

AMAC announces the expansion of its **T-cell receptor (TCR) monoclonal antibody line** (Reader Service No. 105). The specificities of its TCR monoclonals have been confirmed by the polymerase chain reaction on sorted peripheral T cells, clones and TCR transfectants. These monoclonal antibodies cover almost 30 per cent of the human V β repertoire: V β 2, V β 3, V β 5.2, V β 5.3, V β 8, V β 13, V β 17, V β 19, V β 21 and V β 22. Also available are Pan α/β , Pan γ/δ , γ/δ (-V δ 1), V γ 1, V γ 9 and V δ 2. Also provided by AMAC is a T-cell receptor chart describing the TCR structure, genomic organization, genomic recombination and nomenclature for the V β chain.

Hybri-BREScan, the Hybritech **breast cancer marker** is available in an immunoenzymetric and in an immunoradiometric format with the patented Tandem 'sandwich' method (Reader Service No. 106). Hybritech says that it is a powerful marker in detecting a recurrence of breast cancer, particularly in those patients with metastatic disease. Its main clinical application is in monitoring disease progression and response to therapy. A guide for clinicians in a 'Questions and Answers' format provides important information about the assay and its use in monitoring breast cancer patients.

Hycor Biomedical now offers **affinity purified polyclonal antibodies to parathyroid hormone**, both C-terminal (39-84



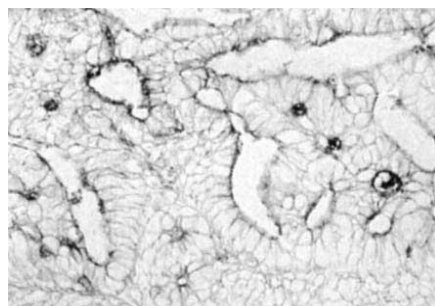
Affinity purified polyclonal antibodies to parathyroid hormone.

segment) and N-terminal (1-34 segment) (Reader Service No. 107). The company says that typical performance using the C-terminal antibody coated to a solid phase and the N-terminal antibody radiolabelled shows a sensitivity (that is, least detectable dose) of less than 1 pg ml^{-1} , a percentage

B/T of about 25 per cent at $1,500 \text{ pg ml}^{-1}$, a maximum binding of 50,000-100,000 pg ml^{-1} , and a high dose hook that is greater than 500,000 pg ml^{-1} . Hycor manufacturers a line of monoclonal, polyclonal and affinity purified antibodies; custom, OEM services and private label are also available.

Detection systems

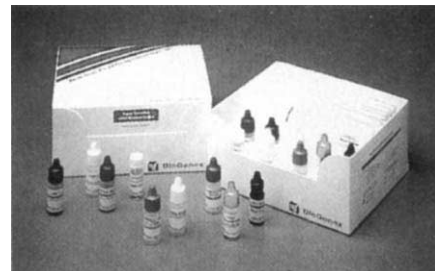
Designed for use with **horseradish peroxidase-based detection systems**, such as the Vectastain Elite ABC kits, the Vector SG substrate produces a bluish-grey reaction



Vector SG substrate and DAB with the Vectastain Elite ABC and anti-CEA and anti-Pan-cytokeratin in human metastatic colon cancer tissue.

product that can be dehydrated and permanently mounted (Reader Service No. 108). Applicable to single-, double- or triple-labelling in immunohistochemistry, Vector Laboratories says that its Vector SG substrate appears to provide particularly intense staining of brain neurons. The product is also useful for the detection of protein, DNA or RNA targets on nitrocellulose or nylon blots. Packaged as concentrated solutions in dropper bottles, the Vector SG substrate working solution can be made simply by dispensing drops of stock reagents into PBS buffer. Vector says that incubation for 2-10 min generally provides good staining intensity in immunohistochemistry. Each kit provides sufficient reagents to stain about 2,000 tissue sections or one hundred and fifty 100 cm^2 blots.

The new **mRNA detection system** from BioGenex Laboratories is optimized for the detection of mRNA in formalin-fixed, paraffin-embedded tissue sections using *in situ* hybridization (Reader Service No. 109). The kit is optimized for use with oligonucleotide probes (20-50 bases long) and contains



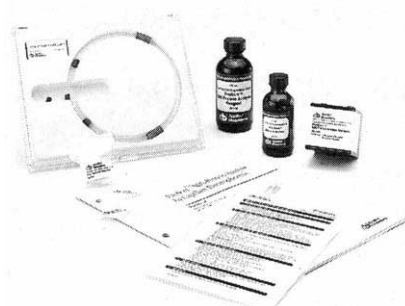
BioGenex's mRNA detection system.

RNase-free reagents that are designed to ensure preservation of the target during the procedure. To provide optimal sensitivity, the kit incorporates a signal amplification system: after hybridization of a biotinylated probe, an anti-biotin antibody is added. Then, following addition of a biotinylated secondary antibody, the slide is incubated with a streptavidin-conjugated alkaline phosphatase label. BioGenex says that this enzyme conjugate, designed specifically for *in situ* hybridization, produces minimal background. This ready-to-use kit can be used with any biotinylated probe, and contains all the detection reagents needed for the hybridization of 50 tissue sections.

The Elite range of **nonradioactive DNA labelling and detection kits** from Cambridge Research Biochemicals is designed to offer an alternative to ^{32}P (*Reader Service No. 110*). Optimized to allow the detection of less than 0.1 pg of target DNA, the system has also been developed to ensure low backgrounds, highly reproducible yields of labelled DNA and reliable stripping and reprobing performance. The DNA is labelled by the incorporation of a hapten using a random priming reaction. Subsequent detection is with a monoclonal antibody-alkaline phosphatase conjugate. The signal is visualized with either chemiluminescent or colour substrates. The Elite starter kit includes sufficient reagents for 15 labelling reactions and detection on 2,000 cm² of membrane.

■ HPLC/electrophoresis

Applied Biosystems has set out to make the **molecular weight determination of proteins** faster and easier with the introduction of its new ProSort SDS-Protein Analysis Kit for use with its electrophoresis systems (*Reader Service No. 111*). ProSort is a polymer solution sieving matrix that enables reproducible molecular weight estimation by capillary electrophoresis, while providing an accuracy and resolution that is comparable to standard SDS-PAGE. ABI says that it offers the advantages of capillary

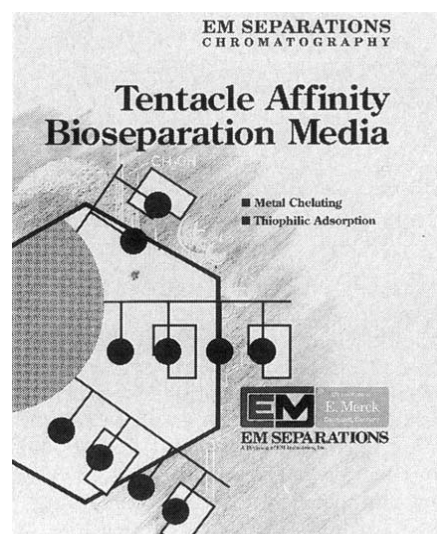


ProSort — ABI's protein analysis kit for capillary electrophoresis.

electrophoresis — fast separation, low sample requirements and built-in quantitation — while eliminating the need for gel preparation, staining and other time-consuming requirements of SDS-PAGE. Sample preparation for ProSort is the same as with conventional SDS-PAGE. However, ABI says that more precise quantitation is possible using the ultraviolet detector built into the capillary electrophoresis system. Proteins ranging from 14 through to 205 kilodaltons can easily be separated in one run. The kit also enables the separation of glycoproteins.

Phenomenex is offering two **HPLC packing materials for the separation of fullerene compounds**, known as Bucky Balls, whose recent discovery has spurred significant interest into their physical and chemical properties (*Reader Service No. 112*). The company's bonded reverse phase material BuckySep-RP, which is said to have a high capacity and good resolving power for fullerenes, can be used with simple solvent systems. Another material, BuckySep-GPC, provides fast and simple analysis and purification of fullerenes via π -electron interaction mechanisms.

EM Separations has made available new literature describing the benefits of its **tentacle affinity bioseparation media** over more conventional affinity supports (*Reader Service No. 113*). Synthesized with metal chelating and thiophilic ligands attached to



New tentacle affinity bioseparation media brochure.

tentacle polymer chains, tentacle affinity bioseparation media is designed to reduce the contact between the biomolecule and the matrix, thus suppressing nonspecific interactions. The tentacle location avoids distortion and eliminates steric interferences, the company says. Other advantages include a higher loading capacity with greater recovery, and better selectivity and more reproducibility for scale-up. The brochure outlines applications for immobilized metal affinity chromatography and thiophilic adsorption affinity chromatography.

■ Catalogues

Microbes and Cells at Work, 2nd Edition is now being offered by the American Type Culture Collection (*Reader Service No. 114*).

The 1993/1994 **enzyme catalogue** is now available from Biozyme Laboratories (*Reader Service 115*). □

These notes are compiled by Diane Gershon from information provided by the manufacturers. For more information, fill in the reader service card bound inside the journal.

ADVERTISEMENTS

\$3.50/Base Custom DNA Synthesis

No Set-up Fee,
Shipped in 24 hours.

**1-800-788-4362
1-800-786-4362 Fax**

**KEYSTONE
Laboratories, Inc.**

Reader Service No.33

Let the experts do the job

ABIMED now offers high quality custom synthesized peptides at very competitive prices. Call or FAX for further information.

- * Quantity 5-50 mg
- * Up to 25 residues
- * HPLC analysis included
- * Crude or purified
- * Mass spectrum available

ABIMED Analysen-Technik GmbH
PB 11 11, D-4018 Langenfeld, FRG. Phone: (+49)-2173-8905-0 FAX: (+49)-2173-8905-77

Reader Service No.6

Highest Quality Oligonucleotides

Synthesised, purified and shipped in 3 days

£2.75 per base

(All inclusive)

0.2 μM scale (approx. 500 μg)

- Antisense (completely or partially sulphurised)
- Extensive range of labelling and modifications
- Full design and analysis facilities
- Gene cassettes, bulk order and contract discounts

Molecular Medicine Unit
King's College School of Medicine & Dentistry
Tel. 071 - 326 3126 FAX 071 - 733 3877

Reader Service No.56

Triplet repeats on the rise

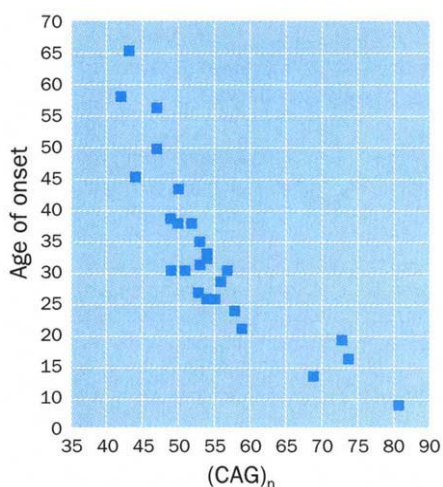
A direct search for expanding trinucleotide repeat sequences has paid off with the identification of the latest disorder caused by this mechanism — spinocerebellar ataxia type 1.

It was only four months ago that the gene responsible for Huntington's disease was cloned¹. Identification of the gene was itself the cause for high excitement, which grew all the more intense with the news that it harbours yet another example of expanding trinucleotide repeat sequences. Such sequences had already been implicated in three other neurological disorders — fragile X syndrome, myotonic dystrophy and spinobulbar muscular atrophy (SBMA)².

In a matter of just a couple of years, triplet repeats have become the most talked about phenomenon in human genetics³. Whether they will be confined to neurological disorders remains to be seen — speculation persists that they may underlie diseases such as breast cancer — but the latest example to join the list is another classic neurological disease. Writing in this month's *Nature Genetics*, Harry Orr, Huda Zoghbi and colleagues show that an expanding CAG repeat is associated with spinocerebellar ataxia type 1 (SCA1)⁴, and the way in which they made the discovery may expedite the search for similar disease genes.

The hereditary spinocerebellar ataxias are a genetically heterogeneous group of disorders, of which SCA1 is a prevalent form. The disease is characterized by neurodegeneration of the cerebellum, spinal cord and brainstem, and, like Huntington's disease, it becomes evident only in adulthood. Death usually results 10–20 years after onset of the disease. The gene for SCA1 was mapped to the short arm of chromosome 6 almost 20 years ago, and as the region containing the gene had been cloned there was little doubt that a search of the megabase or so of DNA using conventional positional cloning techniques would eventually give up the gene.

But faced with a quest that could take years, Orr *et al.* decided to attempt a short



Correlation between expansion of CAG triplet repeat and age-of-onset in spinocerebellar ataxia (SCA1). Note the lack of overlap between repeat lengths of healthy and affected individuals (see ref. 4 for further details).

cut. They reasoned that as SCA1 exhibits anticipation — one of the hallmarks of triplet repeat disorders in which the illness becomes progressively more severe in successive generations, as gauged by earlier onset — it could well be another triplet-repeat disorder and be identifiable by that characteristic. Indeed the effectiveness of such a direct approach had already been demonstrated by Caskey and colleagues in cloning the myotonic dystrophy gene⁵.

So Orr *et al.* screened cosmids from the critical region with oligonucleotide probes designed to detect stretches of triplet repeats, and — lo and behold — isolated a small genomic sequence that hybridized to the CAG-repeat probe. When they examined the size of this repeat in families with SCA1, they found the distinctive pattern of repeat expansion in affected members reminiscent of the CAG expansions in Huntington's disease. In looking at more than 30 individuals with SCA1, Orr *et al.* found that affected members possess about twice the number of repeats (43–81) as do healthy unaffected individuals (25–36). Interestingly, in the cases examined so far, there is no overlap between the top end of the normal spectrum and the lower values in the affected range.

Orr *et al.* also examined the correlation between the number of repeats and the severity of the disease, as indicated by the age-of-onset. Patients with the juvenile-

onset form of SCA1 tended to inherit larger expansions (typically 59–81) than those with later-onset disease (at least 43 repeats) (see figure). Although the authors have not yet sequenced the entire gene that contains the CAG repeats, they do show that the triplet repeat is detected in a ten-kilobase RNA transcript found chiefly in brain and skeletal muscle. Moreover, the repeat is present in an open-reading frame in the genomic sequence. All in all, it looks as if the repeat is translated as a putative polyglutamine stretch in the corresponding protein, as is the case in SBMA and Huntington's disease.

This new discovery begs any number of questions, of which two stand out. First, how many more triplet-repeat disorders are yet to be identified? There could well be quite a few, and the success of at least two groups in searching directly for such repeats^{4,5} means that this is likely to become a popular strategy for tracking down the genes concerned. Two regions that are high on the list for examination in this way are described in a couple of other papers in this month's *Nature Genetics*: two other autosomal dominant ataxias which exhibit variable age-of-onset, SCA type 2 and Machado-Joseph disease, have now been mapped by linkage to the long arm of chromosomes 12 and 14, respectively^{6,7}. Another method which may prove useful is called repeat expansion detection (RED)⁸. This technique uses the polymerase chain reaction to amplify large-scale expanded triplet repeats directly from genomic DNA, and has already been used to identify a variable CTG sequence on chromosome 18 (although the clinical consequences of this expansion are unclear).

The other question concerns the biochemical mechanisms by which trinucleotide expansions result in disease. On that we are almost completely in the dark, and it is where progress is now most sorely needed.

Kevin Davies

Kevin Davies is Editor of Nature Genetics.

1. The Huntington's Disease Collaborative Research Group *Cell* **72**, 971–983 (1993).
2. Mandel, J.-L. *Nature Genet.* **4**, 8–9 (1993).
3. Morel, V. *Science* **260**, 1422–1423 (1993).
4. Orr, H.T. *et al.* *Nature Genet.* **4**, 221–226 (1993).
5. Fu, Y.-H. *et al.* *Science* **255**, 1253–1255 (1992).
6. Gispert, S. *et al.* *Nature Genet.* **4**, 295–299 (1993).
7. Takiyama, Y. *et al.* *Nature Genet.* **4**, 300–304 (1993).
8. Schalling, M., Hudson, T.J., Buetow, K.H. & Housman, D.E. *Nature Genet.* **4**, 135–139 (1993).

Also in this month's *Nature Genetics*: sequence characterization of 3,400 complementary DNAs expressed in human brain; identification of the *Little* murine mutation; mammalian recombination genes homologous to *recA* and *RAD51*; defects in rhodopsin giving rise to stationary night blindness; a ribosomal RNA mutation causing non-syndromic deafness; and the healing of telomeres.